ESTIMATION OF *E. COLI* CONCENTRATIONS FROM NON POINT SOURCES USING GIS

A Thesis

by

KYNA ELLEN MCKEE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2011

Major Subject: Biological and Agricultural Engineering

Estimation of *E. coli* Concentrations from Non Point Sources Using GIS

ESTIMATION OF *E. COLI* CONCENTRATIONS FROM NON POINT SOURCES USING GIS

A Thesis

by

KYNA ELLEN MCKEE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Co-Chairs of Committee,	Raghupathy Karthikeyan
	Patricia Smith
Committee Member,	Raghavan Srinivasan
Head of Department,	Stephen Searcy

August 2011

Major Subject: Biological and Agricultural Engineering

ABSTRACT

Estimation of *E. coli* Concentrations from Non Point Sources Using GIS. (August 2011)

Kyna Ellen McKee, B.S., Texas A&M University

Co-Chairs of Advisory Committee: Dr. Raghupathy Karthikeyan
Dr. Patricia Smith

When developing a Watershed Protection Plan (WPP) or a Total Maximum Daily Load (TMDL), it is often difficult to accurately assess the pollutant load for a watershed because not enough water quality monitoring data are available. According to the Texas Commission on Environmental Quality (TCEQ), there are 274 bacteria impairments in Texas water bodies out of 386 impaired water bodies. Bacteria water quality data are often more sparse than other types of water quality data, which hinders the development of WPPs or TMDLs. The Spatially Explicit Load Enrichment Calculation Tool (SELECT) was used to develop watershed protection plans for four rural watersheds in Texas that are impaired due to *E. coli* bacteria. SELECT is an automated Geographical Information System (GIS) tool that can assess pathogen loads in watersheds using spatial factors such as land use, population density, and soil type. WPPs were developed for four rural Texas watersheds: Buck Creek, Lampasas River, five sub watersheds of the Little Brazos River, and Geronimo Creek. A spatial watershed model was developed to simulate bacteria concentrations in streams resulting from non point sources using SELECT combined with a simple rainfall-runoff model and applied to the Geronimo Creek watershed. The watershed model applies a rainfall-driven loading function to the

potential *E. coli* loads calculated by the output of SELECT. The simulated runoff volumes and *E. coli* concentrations from the model were compared to actual monthly *E. coli* data collected at two sampling sites near the outlet of a subwatershed.

The results show how SELECT methodology was applied to each watershed and adapted based on stakeholder concerns and data availability. The highest potential contributors were identified and areas of concern were highlighted to more effectively apply best management practices (BMPs). The runoff volumes were predicted with very good agreement (E = 0.95, RSR = 0.21 to 0.22) for both sampling sites. The predicted *E. coli* concentrations did not agree with measured concentrations for both sites using eight different methods. The results indicate that the model does not include significant factors contributing to the transport of *E. coli* bacteria but can be modified to include these factors.

ACKNOWLEDGEMENTS

I would like to thank my committee co-chairs, Dr. Karthikeyan and Dr. Smith, and my committee member, Dr. Srinivasan, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. I also want to extend my gratitude to the Texas Water Resources Institute and the United States Geological Survey for partially funding this project with the USGS Water Resources Grant. I would like to thank the Spatial Sciences Laboratory at Texas A&M University for providing data support.

Finally, thanks to my mother, father, and brother for their encouragement and to my fiancé for his love and support.

TABLE OF CONTENTS

Page

ABSTRACT iii

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF FIGURES viii

LIST OF TABLES xi

CHAPTER

I INTRODUCTION 1

1.1 Introduction 1
1.2 Objectives 6

II APPLYING SELECT METHODOLOGY TO THE BUCK CREEK, LITTLE BRAZOS RIVER, AND LAMPASAS RIVER WATERSHEDS 8

2.1 Introduction 8
2.2 Study Areas 9
2.3 Methodology 11
2.4 Results and Discussion 25
2.5 Conclusions 33

III ESTIMATING *E.COLI* CONCENTRATIONS FROM NON POINT SOURCES FOR GERONIMO CREEK 34

3.1 Introduction 34
3.2 Study Area 36
3.3 Methodology 38
3.4 Results and Discussion 61
3.5 Conclusions 80

IV CONCLUSIONS 82

Page

4.1 Conclusions 82
4.2 Recommendations 82

REFERENCES 84

VITA 90

LIST OF FIGURES

FIGURE		Page
2.1	Spatial locations of Buck Creek, Little Brazos River, and Lampasas River watersheds in Texas	9
2.2	Buck Creek watershed land use	15
2.3	Land use of Little Brazos River watershed five tributary watersheds	18
2.4	Lampasas River watershed land use	21
2.5	WMAs area locations in the Lampasas River watershed with deer population density estimations	24
2.6	Total daily potential *E. coli* load resulting from various sources in the Buck Creek watershed	26
2.7	Total daily potential *E. coli* load resulting from various sources in the five tributary watersheds of the Little Brazos River watershed	30
2.8	Total daily potential *E. coli* load resulting from various sources in Lampasas River watershed	32
3.1	Geronimo Creek watershed study area with NCDC rain gauges and water quality sampling sites	38
3.2	Geronimo Creek watershed land use classification	41
3.3	Curve number grid	42
3.4	Flow chart illustrating the calculation of accumulated runoff volume	45
3.5	Total potential *E. coli* load calculated using SELECT for the Geronimo Creek watershed	50
3.6	Flow chart illustrating the calculation of the contributing *E. coli* load	51
3.7	Haberle Road FDC developed using instantaneous flows	53
3.8	SH 123 FDC developed from instantaneous flows	54

FIGURE Page

3.9 Haberle Road FDC developed from SWAT simulated flows 54

3.10 SH 123 FDC developed from SWAT simulated flows 55

3.11 Predicted runoff volumes compared to observed runoff volumes using the best performing method for SH 123 for estimating runoff volumes using a time and baseflow 67

3.12 Predicted runoff volumes compared to observed runoff volumes for the Haberle Road sampling site using the best performing method for estimating runoff volumes using a time and baseflow 68

3.13 Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for SH 123 using instantaneous observed flow delivery factor and the simulated runoff volume 73

3.14 Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for SH 123 using simulated flow delivery factor and simulated runoff volume 73

3.15 Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for SH 123 using instantaneous observed flow delivery factor and observed runoff volume 74

3.16 Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for SH 123 using simulated flow delivery factor and observed runoff volume 74

3.17 Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for Haberle Road using instantaneous observed flow delivery factor and simulated runoff volume 76

3.18 Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for Haberle Road using simulated flow delivery factor and simulated runoff volume..... 76

FIGURE Page

3.19 Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for Haberle Road using instantaneous observed flow delivery factor and observed runoff volume 77

3.20 Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for Haberle Road using simulated flow delivery factor and observed runoff volume..... 77

LIST OF TABLES

TABLE		Page
2.1	Calculation of potential *E. coli* loads from various sources	14
2.2	Source specific potential *E. coli* load ranges for the Buck Creek watershed	25
2.3	Source specific potential *E. coli* load ranges for the five tributaries of the Little Brazos River watershed	28
2.4	Source specific potential *E. coli* load ranges for the Lampasas River watershed	31
3.1	NRCS curve number lookup table	43
3.2	Various times used to estimate runoff volumes in the Geronimo Creek watershed	57
3.3	Different delivery factors used for *E. coli* concentration calculation	58
3.4	Model performance for predicting runoff volume at SH 123	64
3.5	Model performance for estimating runoff volume at Haberle Road	66
3.6	Model performance for *E. coli* concentrations at SH 123	70
3.7	Model performance for estimating *E. coli* concentrations at Haberle Road	71

CHAPTER I

INTRODUCTION

1.1. Introduction

Bacteria are the most common reason for impairment of Texas water bodies (TCEQ, 2008). According to the Texas Commission on Environmental Quality (TCEQ), there are 274 bacteria impairments in Texas water bodies out of 386 impaired water bodies (TCEQ, 2008). Bacterial impairment is usually assessed by measuring the actual concentra+tion of an indicator organism. When the concentration of the indicator organism exceeds the regulatory standards, the stream is considered impaired due to fecal contamination. In the State of Texas, *E. coli* is considered an indicator organism of fecal contamination. While addressing the issue of bacteria impairment in Texas water bodies, a need has emerged for an accurate and simple model that would simulate bacterial transport when developing a Watershed Protection Plan (WPP) or Total Maximum Daily Load (TMDL).

Implementing and developing a TMDL project is costly. According to the United States Environmental Protection Agency (USEPA), "the national average cost of developing TMDLs per water body is estimated to be about $52,000, but can typically range from under $26,000 to over $500,000 depending on the number of TMDLs, their level of difficulty and the extent to which impaired waters are clustered together for

---

This thesis follows the style of Journal of Environmental Management.

TMDL development" (USEPA, 2001b). A considerable amount of time and money is spent while developing a TMDL to allocate pollutant load and to identify potential sources. Usually, TMDL development is done by using extensive water quality models that require a significant amount of resources and time.

Rainfall is an important driving factor in watershed hydrology and water quality (Haan et al., 1994). Bacteria fate and transport is believed to be considerably influenced by surface runoff resulting from rainfall (Haan et al., 1994). Curriero et al. (2001) estimated more than half of the waterborne disease outbreaks were preceded by precipitation events in the last 50 years in the United States. They also found that extreme wet-weather events play an important role in microbial fate and transport (Curriero et al., 2001). The subsequent runoff from rainfall events carries pathogens from the surrounding land surfaces into the streams, causing outbreaks of waterborne diseases such as giardiasis, diarrhea, cryptosporidiosis, etc.

Accurately including rainfall as a parameter in a model is difficult because rain gauges are only able to precisely measure the rainfall at the point the rain gauge is located and not over an entire field or watershed. Rainfall can be spatially distributed over an area using multiple point rain gauges and spatial interpolation with GIS. Spatial interpolation can be done with many different methods. The most popular methods used in distributed hydrologic models include the Nearest Neighborhood (NN) and Inverse Distance Weighted (IDW) methods (Zhang and Srinivasan, 2009). Schuurmans and Bierkens (2006) illustrated different spatial distribution techniques and the applications

to a hydrological model. The results show that using a single rain gauge can produce inaccurate results and differences in the spatial distribution of rainfall can considerably affect the model results.

Models such as Soil and Water Assessment Tool (SWAT) and Hydrological Simulation Program- FORTRAN (HSPF) have been used for modeling bacteria transport. Other simplistic microbial models such as, the potential non point pollution index (PNPI) and a Spatially Explicit Delivery MODel (SEDMOD), and Spatially Explicit Load Enrichment Calculation Tool (SELECT), have been developed to rank the potential pollution impacts of areas from nonpoint sources primarily utilizing land use, geomorphology, and potential sources in the watershed(Fraser et al., 1998; Munafo et al., 2005; Teague et al., 2009).

One commonly used watershed-scale model is SWAT, which simulates long term sediment yield and hydrologic processes on a daily time step. SWAT was developed for use in large ungauged basins and is intended as a long term planning tool to predict the impacts of management on water, nutrient, and sediment yield (Arnold et al., 1998). SWAT utilizes geographically referenced data for input parameters (Parajuli et al., 2009). A microbial sub-model was incorporated in SWAT 2005 (Sadeghi and Arnold, 2002). It contains functional relationships for bacteria die-off and regrowth rates that can cover a range of pathogenic bacteria. The model allows for risk evaluation associated with agricultural practices of nutrients, pathogens, and sediment loadings (Sadeghi and Arnold, 2002).

The SWAT microbial sub-model requires the model to be calibrated to improve the model performance. Benham et al. (2006) calibrated SWAT for bacterial contamination in a watershed using more than two years of daily flow values. Manual calibration was performed by adjusting one parameter at a time and comparing simulated and observed hydrographs (Benham et al., 2006). Model calibration was also necessary in a study done by Parajuli et al. (2009) because the SWAT default parameters had low model efficiency for daily flow (Parajuli et al., 2009). In small Texas watersheds, data are frequently not available for model calibration. The historical data collected are often sporadic and more frequent monitoring data for an adequate time period is not available until a WWP or TMDL project is ending. While developing TMDLs, the model is often run before collecting additional monitoring data.

Another frequently used hydrologic model is HSPF, made available through the U.S. Geological Survey (USGS). It is able to simulate hydrologic and water quality processes on pervious and impervious surfaces, in streams, and well-mixed impoundments for extended periods of time (USGS, 2010). Bicknell et al. (2001) describes the use of HSPF, to simulate hydrologic processes and the related water quality constituents on pervious and impervious land surfaces in streams. The input data required to run HSPF for watershed water quality simulation include meteorologic records of precipitation, estimates of potential evapotranspiration, air temperature, wind, humidity, point sources and other physical measurements (USGS, 2010).

HSPF requires extensive monitoring data as input data for model calibration because the model relies on empirical relationships for calculations (Borah and Bera,

2004). HSPEXP, the expert system for the calibration of HSPF, interactively allows the user to change model parameters to optimize calibration (Paul et al., 2004). Paul et al. (2004) calibrated HSPF using historical daily mean stream flow data from a USGS gauge station for the simulation period. The input data required to run and calibrate HSPF effectively requires the utlization of a USGS gauge station located in the watershed. This is problematic when the model is needed to represent ungauged watersheds such as many of the impaired watersheds in Texas. Surrogate parameters from a similar watershed must be used to define the hydrology of the watershed in order to run the model for an ungauged watershed (USGS, 2010).

The SELECT methodology was developed to characterize *E. coli* sources from point and non-point pollution in watersheds where bacterial contamination is a concern for WPP or TMDL development (Teague et al., 2009). Automated SELECT provides a graphical user interface (GUI) within ArcGIS 9.X. Project parameters can be adjusted by the user for pollutant loading scenarios specific to a watershed. SELECT simulates potential bacterial loading by source and area characterization for different management scenarios (Riebschleager et al., 2011).

The pollutant connectivity factor (PCF) is another component of SELECT. The PCF utilizes the potential total pollution resulting from SELECT and weighs the influence of driving forces of contamination. The results from the PCF are a ranking of the potential contribution which can be used to show the areas in the watershed that are vulnerable to contributing bacteria to water bodies (Riebschleager et al., 2011). The PCF is only able to indicate “hot spots” within the watershed and does not include

complete fate and transport processes in the watershed. SELECT is able to estimate potential bacterial loadings and indicate vulnerable areas but is unable to determine actual bacterial loads in the water bodies. SELECT can be combined with a fate and transport model to estimate the actual bacterial loadings in the water bodies.

In ungauged watersheds, historical bacteria data is sparsely available. It is expensive to collect monitoring data. The USEPA estimates a cost of approximately $17 million a year for water quality monitoring to support the development of all national TMDL projects (USEPA, 2001b). Current bacteria models require extensive monitoring data within the watershed for calibration or they cannot predict actual *E. coli* concentrations in the water body. A simple model that predicts actual bacteria concentrations in a water body is needed in order to develop TMDLs or WWPs within the State of Texas. Also, this model should incorporate stakeholder inputs while developing TMDLs and WPPs. Typically the stakeholder group consists of farmers, ranchers, common public, administrators, and extension personnel living in the watersheds.

1.2. Objectives

The overall objective of this research project was to develop a conceptual model in ArcGIS 9.3 utilizing the potential *E. coli* load estimated by SELECT to simulate *E. coli* concentrations occurring in a rural Texas stream. It was presumed that precipitation is the main driving factor for the transport of *E. coli* bacteria from sources to the stream. Also the affects of temperature were negligible, since in Texas watersheds the monthly normal daily mean temperatures do not vary from month to month by more than 10 °F.

The specific objectives were:

(1) To apply the SELECT methodology for different watersheds in Texas incorporating stakeholder inputs;

(2) To apply SELECT to the Geronimo Creek watershed using stakeholder inputs concerning the *E. coli* sources and the population densities;

(3) To develop an automated rainfall-runoff model in ArcGIS 9.3 utilizing rain gauges located in and around the Geronimo Creek watershed and to estimate the *E. coli* concentrations in the creek.

CHAPTER II

APPLYING SELECT METHODOLOGY TO THE BUCK CREEK, LITTLE BRAZOS RIVER, AND LAMPASAS RIVER WATERSHEDS

2.1. Introduction

Accurately assessing the pollutant load for a watershed, for the development of a Total Maximum Daily Load (TMDL) or Watershed Protection Plan (WPP) is difficult because generally not enough water quality monitoring data is available. A WPP is a stakeholder driven process to restore or protect the water quality of a specific water body. The most common reason for the impairment of waterbodies in Texas and across the United States is bacteria (TCEQ, 2008; USEPA, 2008). Out of a total of 386 impaired water bodies in Texas, 274 are impaired due to bacteria (TCEQ, 2008). The development of bacteria WPPs or TMDLs can be hindered due to the spare availability of bacteria water quality data.

SELECT is an automated Geographic Information System (GIS) tool that can be applied to assess potential *E. coli* loads in a watershed based on spatial factors such as land use, population density, and soil type (Teague, et al., 2009). SELECT is able to calculate a potential *E. coli* load and highlight areas of concern for best management practices (BMPs) to be implemented. The potential *E. coli* load in SELECT is calculated by distributing the contributing sources spatially over the entire watershed. When applying SELECT, the population densities of potential contributors are determined with stakeholder input to accurately represent the watershed. However, it

should be noted that potential *E. coli* loads generated using SELECT are the worst case scenario because the tool the largest amount of contribution possible from individual sources.

2.2. Study Areas

Three impaired watersheds in Texas, the Little Brazos River watershed, the Buck Creek watershed, and the Lampasas River watershed, were selected to apply SELECT methodology to predict potential *E. coli* loads resulting from various sources.

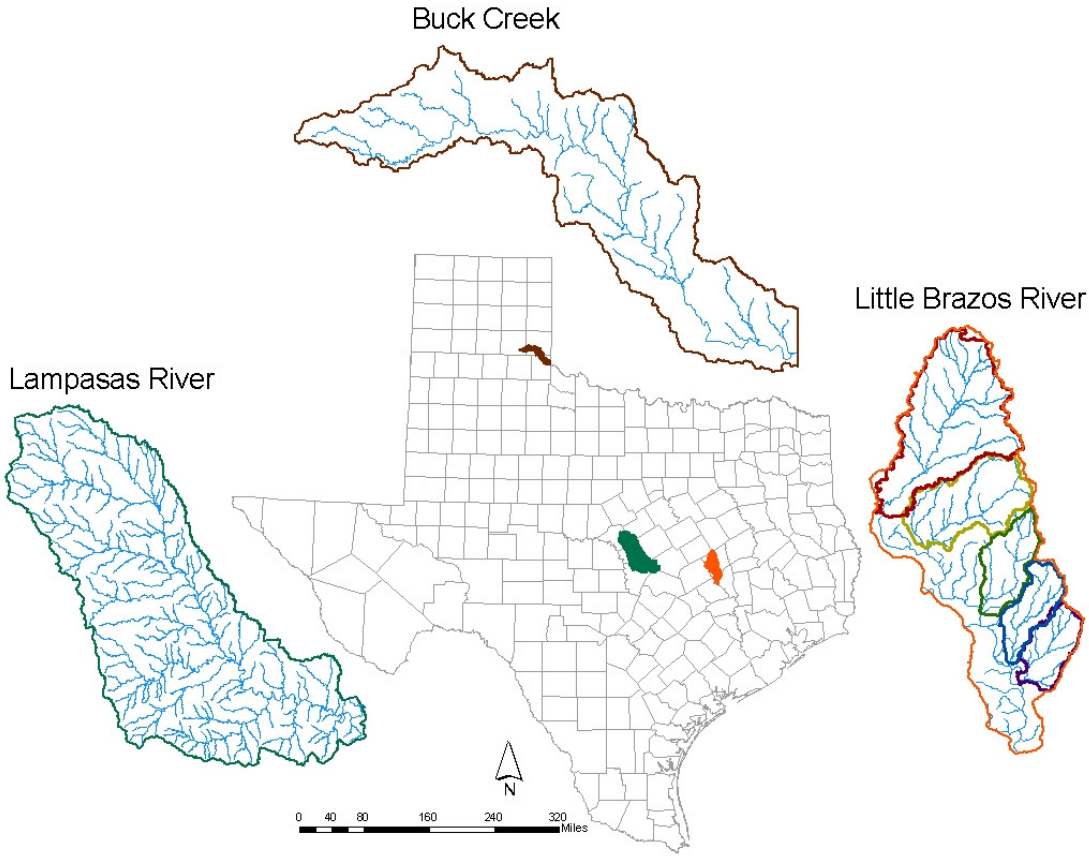

Figure 2.1. Spatial locations of Buck Creek, Little Brazos River, and Lampasas River watersheds in Texas.

2.2.1. Buck Creek Watershed

Buck Creek (Figure 2.1) is a small, unclassified stream that originates southwest of Hedley, Texas in Donley County and flows 109 kilometers (68 miles) across the Oklahoma border to its confluence with the Prairie Dog Town Fork of the Red River. Buck Creek is classified as an impaired water body due to bacterial contamination under the EPA Clean Rivers Act 303 (d). The study area only includes the portion of the watershed located in Texas which encompasses an area of 74,851 hectares (184,960 acres). Buck Creek is located across Donley County, Childress County, and Collingsworth County in the Texas Panhandle. The watershed is mostly agricultural populated with a few rural towns.

2.2.2. Little Brazos River Watershed

The Little Brazos River watershed (Figure 2.1) located in the central Brazos River basin consists of one classified water body. This watershed contains five tributaries impaired for bacteria; located within very close proximity of each other in Robertson County, the subwatersheds share similar land use and water quality characteristics. The five impaired tributaries of the Little Brazos River watershed are Campbells Creek, Mud Creek, Pin Oak Creek, Spring Creek, and Walnut Creek. The watershed area containing the watersheds of the tributaries encompasses 84,693 hectares (209,280 acres) that lies almost entirely within Robertson County. The land use in the area is primarily agricultural consisting of range and pastureland with mixed areas of forested lands and several small towns and communities.

2.2.3. Lampasas River Watershed

The Lampasas River watershed is located in south central Texas and begins in Hamilton County and flows 121 kilometers (75 miles) through Lampasas, Burnet, and Bell Counties. The study area only includes the length of the Lampasas River until it is dammed and forms Stillhouse Hollow Lake. Including tributaries the Lampasas River watershed encompasses 322,320 hectares (796,469 acres). The land use for the Lampasas River watershed is primarily agricultural containing rural towns with the exception of the lower portion of the watershed which contains a portion of the Fort Hood-Killeen area.

2.3. Methodology

The Spatially Explicit Load Enrichment Calculation Tool (SELECT) methodology developed by Biological and Agricultural Engineering Department and Spatial Sciences Laboratory at Texas A&M University was used to independently characterize potential *E. coli* sources and estimate daily potential *E. coli* loads for the Buck Creek watershed, the Lampasas River watershed, and each of the five Little Brazos River tributary watersheds. SELECT is an analytical approach for developing an inventory of potential bacterial sources, particularly nonpoint source contributors, and distributing their potential bacterial loads based on land use and geographical location.

A thorough understanding of the watershed and potential contributors that exist is necessary to estimate and assess bacterial load inputs. Land use classification data and data from state agencies, municipal sources, and local stakeholders on the number and distribution of pollution sources are used as inputs in a Geographical Information

Systems (GIS) software format. The watershed is divided into multiple smaller subwatersheds based on elevation changes along tributaries and the main segment of the water body. Pollutant sources in the landscape can then be identified and targeted where they are most likely to have significant effects on water quality, rather than looking at contributions on a whole-watershed basis. Typically, a stakeholder group consists of farmers, ranchers, common public, administrators, and extension personnel living in the watersheds. The role of a stakeholder group when applying SELECT to a watershed is to review inputs into SELECT. Individual stakeholders apply personal knowledge of the watershed to make those inputs as accurate as possible.

The land use was verified by stakeholders and it was suggested that the land use categorized as crop land should be categorized as managed pasture for the Little Brazos River watershed and an additional land use type was added to crop land for the Buck Creek and Lampasas River watersheds. Visual outputs of the program allow a decision maker or stakeholder to easily identify areas of a watershed with the greatest potential for contamination contribution and enable the decision maker to use that information to help formulate management strategies to include in a WPP or TMDL implementation plan.

2.3.1 Potential *E. coli* Load Estimation

Sources potentially contributing to a watershed are determined by stakeholders. Identified sources were: cattle, goats, sheep, horses, confined animal feeding operations (CAFOs), poultry operations, deer, feral hogs, dogs, on-site wastewater treatment systems (OWTSs), and wastewater treatment facilities (WWTFs). The analysis was conducted at a 30 meter by 30 meter spatial resolution. For each source, it is first distributed to the suitable areas in the watershed and then the *E. coli* load was calculated using the equations in Table 2.1. The fecal production rates for the sources were calculated using the highest in the range of values in the EPA guidance (USEPA, 2001a) for all of the *E. coli* sources. After the potential *E. coli* loads are calculated, the results are aggregated at the subwatershed level to easily distinguish areas of concern.

2.3.2. Potential *E. coli* Sources in the Buck Creek Watershed

Cattle, feral hogs, and deer were identified as potential fecal contributors in the Buck Creek watershed. These were determined to be potential fecal contributors by state agencies, stakeholders, and the capabilities of the model.

Table 2.1. Calculation of potential *E. coli* loads from various sources.

Source	***E. coli* Load Calculation**
Cattle	$EC = \#\ Cattle * 10 * 10^{10}\ cfu/day * 0.5$
Horses	$EC = \#\ Horses * 4.2 * 10^{8}\ cfu/day * 0.5$
Sheep and goats	$EC = \#\ Sheep * 1.2 * 10^{10}\ cfu/day * 0.5$
Confined Animal Feeding Operations	$EC = \#\ Permitted\ Head * 10 * 10^{10}\ cfu/day * 0.2 * 0.5$
Poultry Operations	$EC = Maximum\ Amount\ of\ Litter\ Utilized\ On\text{-}Site * 44{,}000\ cfu/gram$
Deer	$EC = \#\ Deer * 3.5 * 10^{8}\ cfu/day * 0.5$
Feral hogs	$EC = \#\ Hogs * 1.1 * 10^{9}\ cfu/day * 0.5$
Dogs	$EC = \#\ Households * \frac{1\ dog}{Household} * 5 * 10^{9} cfu/day * 0.5$
On-site Wastewater Treatment Systems	$EC = \#OWTSs * Failure\ Rate * \frac{10 * 10^{6} cfu}{100\ mL} * \frac{70\ gal}{\frac{person}{day}} * \frac{Avg\ \#}{Household} * \frac{3758.2\ mL}{gal} * 0.5$
Wastewater Treatment Facilities	$EC = Permitted\ MGD * \frac{126\ cfu}{100\ mL} * \frac{10^{6}\ gal}{MGD} * \frac{3758.2\ mL}{gal}$

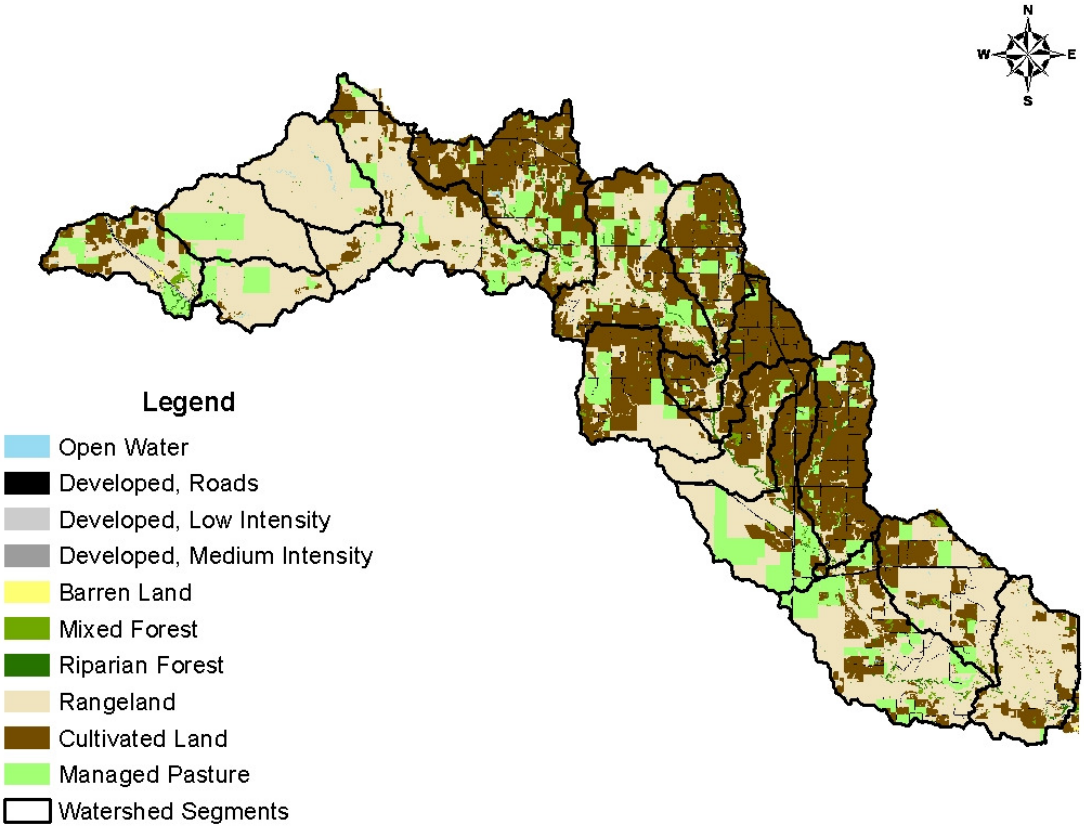

Figure 2.2. Buck Creek watershed land use.

2.3.2.1 Cattle

Populations of cattle in the Buck Creek watershed consist of those grazed on rangeland and those grazed on managed pasture (Figure 2.2). Using an average NRCS stocking rate for rangeland of 10 ha/animal (25 ac/animal) and for managed pasture of 3 ha/animal (8 ac/animal) in Childress, Collingsworth and Donley counties, the total watershed population of cattle is currently estimated at 6,640 head (454 kg animal units). Rangeland cattle accounted for 3,664 head and were evenly distributed in the rangeland,

mixed forest, and riparian forest land uses, (Figure 2.2) while the remaining (2,976) managed pasture cattle were evenly distributed in the managed pasture land use. These cattle numbers and distributions were verified with watershed stakeholders and determined to be representative of the Buck Creek watershed. The potential *E. coli* loads were calculated (Table 2.1) separately for range and pasture cattle and added together to create the total potential *E. coli* loads resulting from cattle.

2.3.2.2 Deer

Deer populations estimated in Buck Creek comprised of white-tailed and mule deer. The SELECT methodology is not able to distinguish between separate deer species, therefore combining the two populations into one was the most feasible scenario. The Texas Parks and Wildlife Department (TPWD) study conducted by Lockwood (2005) provided initial population estimates and associated animal densities for areas as near to Buck Creek as possible. Using this information as a starting point, stakeholders were asked to provide input on the size and distribution of the deer herds in the watershed. In total, approximately 5,143 deer (990 Mule deer and 4,153 White-tails) were assumed to reside in the watershed and were applied over contiguous areas of rangeland, managed pasture, mixed forest, riparian forest and cultivated land uses (Figure 2.2) at an average rate of 15 hectares (36 acres) per animal. Using the equation from Table 2.1, daily potential *E. coli* loads resulting from deer were estimated.

2.3.2.3 Feral Hogs

No accurate estimate of feral hog numbers in the Buck Creek watershed exists. All stakeholders were asked to provide input regarding feral hog numbers in Buck Creek.Using this feedback, an acceptable population estimation of 7,310 animals was determined. Stakeholders also indicated that the feral hog population should be distributed across rangeland, barren land, managed pasture, cultivated land, mixed forest, and riparian forest land uses (Figure 2.2) within a 100 meter buffer around streams. Applying this population estimate to these land uses resulted in a population density of 10 hectares (25 acres per animal) for the entire watershed area. Then, daily potential *E. coli* loads resulting from feral hogs were estimated (Table 2.1).

2.3.3. Little Brazos River Watershed Potential *E. coli* Sources

The following potential *E. coli* sources were considered in estimating total potential *E. coli* loads resulting from each subwatershed. To simplify the modeling purposes, the stocking rates for livestock, wildlife, and feral hogs were consistently applied for all five subwatersheds.

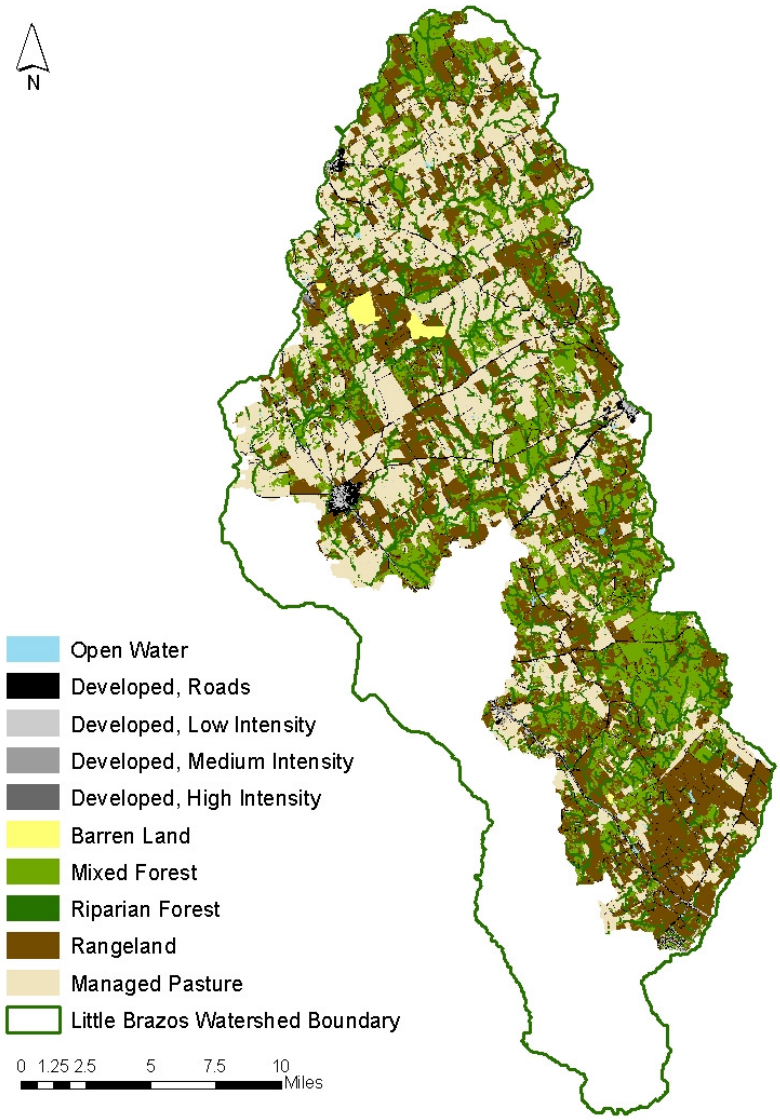

Figure 2.3. Land use of Little Brazos River watershed five tributary watersheds.

2.3.3.1. Livestock –Cattle

The cattle population was calculated as two separate management practices, pasture cattle and range cattle to account for the different stocking rates associated with the different types of cattle management. For pasture cattle, the stocking rate of 0.8

hectares (two acres) per animal unit was applied uniformly over the managed pasture lands (Figure 2.3) in each subwatershed. Daily potential *E. coli* loads resulting from pasture cattle were estimated using the equation from Table 2.1. For range cattle, the stocking rate of two hectares (five acres) per animal unit was applied uniformly over the rangeland, mixed forest, and riparian forest (Figure 2.3) in each subwatershed. The pasture cattle and range cattle results were then added together spatially to create the resulting potential loadings from cattle for each watershed.

2.3.3.2. Wildlife - Deer

For deer, a density of 15 hectares (37 acres) per animal unit was applied over contiguous areas of the rangelands, managed pasture lands, mixed forest, and riparian forests (Figure 2.3) in each subwatershed. The number of deer estimated using this density and the equation from Table 2.1 were used to calculate the daily potential *E. coli* load resulting from deer.

2.3.3.3. Feral Hogs

For feral hogs, a density of 8 hectares (20 acres) per animal unit was applied uniformly across range lands, managed pasture lands, mixed forest, and riparian forests (Figure 2.3) within a 100 m buffer around the stream network of each subwatershed. Daily potential *E. coli* loads resulting from feral hogs were estimated using the density and the equation from Table 2.1.

2.3.3.4. Poultry Operations

For poultry operations, the maximum litter to be utilized on-site in tons per day was applied uniformly over the subwatershed where the poultry operation is located.

Since poultry litter is not applied consistently throughout the year and is applied only once annually, the *E. coli* load calculated is for the day that the litter is applied. This demonstrates the worst case scenario in the watershed during that particular day. The *E. coli* concentration used was 44,000 CFU per gram of broiler litter. Using the maximum litter to be utilized on-site and *E. coli* concentration in broiler litter, potential *E. coli* loads resulting from poultry litter application on one particular day were estimated.

2.3.3.5. On-site Wastewater Treatment Systems (OWTSs)

For on-site wastewater treatment systems (OWTSs), the *E. coli* load was calculated using the formula from Table 2.1. The number of systems was the number of homes from the 2000 Census Blocks with the homes removed from areas falling within urban areas. The failure rate was calculated from the Septic Drainfield Limitation Class using the SSURGO soil database. The failure rate for each limitation class is as follows: very limited as 15%, somewhat limited as 10%, slightly limited as 5%, and not rated as 15%. The people per home were the average household size from the 2000 census blocks. This resulted in daily potential *E. coli* load resulting from septic systems.

2.3.3.6. Wastewater Treatment Facilities (WWTFs)

For wastewater treatment facilities (WWTFs), the maximum permitted discharge rate and the *E. coli* concentration of 126 CFU/100 mL (Table 2.1) was applied to the subwatershed in which the WWTFs were located. There were three WWTFs located in the Little Brazos Watersheds, two located in the Mud Creek Watershed and one located in the Walnut Creek Watershed.

2.3.4 Lampasas River Watershed Potential *E. coli* Sources

To estimate potential *E. coli* loads in the Lampasas River Watershed, domestic, livestock, and wildlife sources were considered and distributed on appropriate land use (Figure 2.4). Potential domestic contributors included OWTSs, dogs, and WWTFs. Livestock included horses, goats, sheep, cattle, and confined animal feeding operations (CAFOs). Deer and feral hogs were identified as the wildlife contributing to the contamination that could be feasibly modeled.

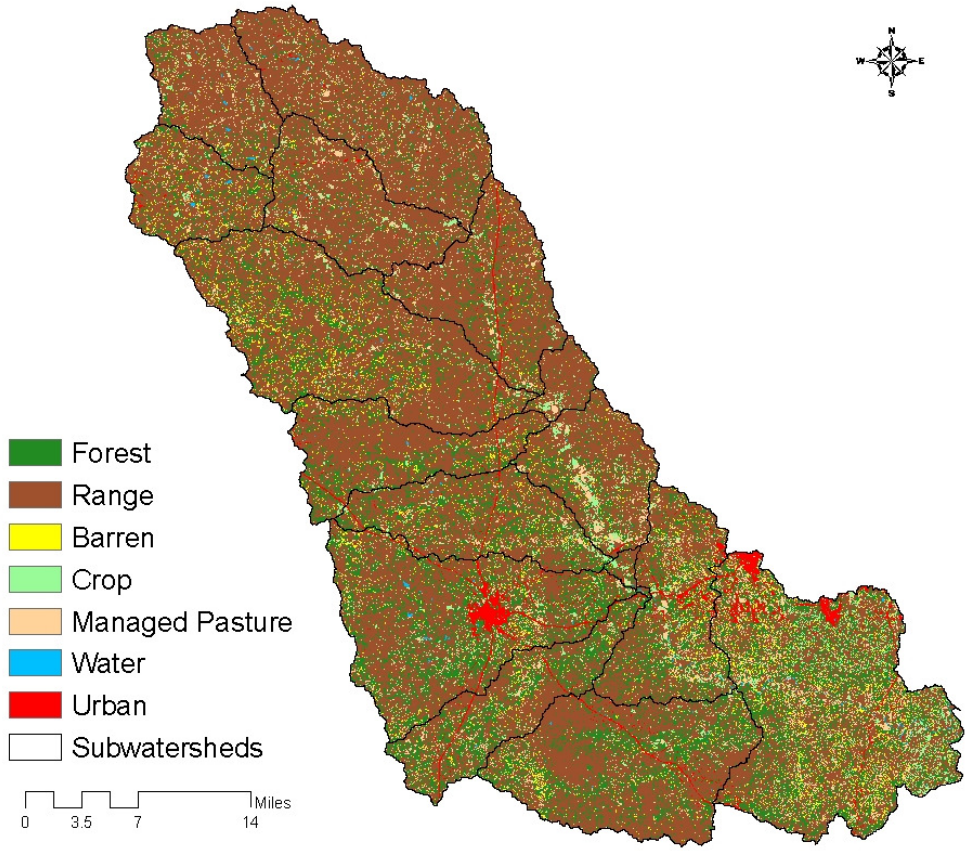

Figure 2.4. Lampasas River watershed land use.

2.3.4.1. On-site Wastewater Treatment Systems

The potential *E. coli* load for OWTSs was calculated using the equation from Table 2.1. For OWTSs, spatially distributed point data of each household was collected from residential 911 address data and households within Certificate of Convenience and Necessity (CCN) areas were removed to not include households being serviced by a WWTF. The people per home were the average household size from the 2000 census blocks. A constant discharge of 265 liters (70 gallons) per person per day was used in the calculations. A failure rate was determined for the OWTSs using SSURGO soil limitation classes to calculate the percentage of *E. coli* contributing to the watershed due to septic failure.

2.3.4.2. Dogs

The potential *E. coli* load resulting from dogs was calculated using the equation from Table 2.1. A density of one dog per household was applied to the residential 911 addresses resulting in an estimated dog population of 10,775.

2.3.4.3. Wastewater Treatment Facilities

The Lampasas River watershed contained two WWTFs located in separate subwatersheds. For wastewater treatment facilities (WWTFs), the maximum permitted discharge and the *E. coli* concentration of 126 CFU/100 mL was applied to the subwatershed in which the WWTFs were located.

2.3.4.4. Livestock

The population for livestock was estimated using the 2007 Census of Agriculture by considering only the number of animals located in the watershed for each county.

The percentage of the watershed located in each county was calculated and that percentage was used to determine the number of animals in the watershed for each county from the total county population. Goats, sheep, and cattle were evenly distributed amongst the range, forest, and managed pasture land uses (Figure 2.4) and had estimated populations of 11,162; 7,311;and 34,338 respectively, for the entire watershed area (USDA-NASS, 2007). Horses were evenly distributed on rangelands (Figure 2.4) and had an estimated population of 1,288 animals (USDA-NASS, 2007).

2.3.4.5. Confined Animal Feeding Operations (CAFOs)

Three confined animal feeding operations (CAFOs) were located in the Lampasas River watershed. For CAFOs, the permitted number of head of cattle was used to determine the potential *E. coli* load for the subwatershed where the CAFOs are located. The *E. coli* production rate of 10^{11} CFU per animal per day was applied with an assumed treatment efficiency of 80% resulting in an *E. coli* load of 2×10^{10} CFU per animal being applied.

2.3.4.6. Deer

Wildlife management associations (WMAs) are located in areas around the Lampasas River watershed shown in Figure 2.5 and have population density estimations for deer located in these specific areas. The deer densities within the WMAs were applied uniformly over the entire area of the WMA without considering land use types. For the areas not within a WMA, a density of 100 deer per 405 hectares (1000 acres) was applied over the entire area of the watershed without considering land use types. An

estimated population of 84,739 deer was used with the equation from Table 2.1 to estimate the potential *E. coli* load resulting from deer for the watershed.

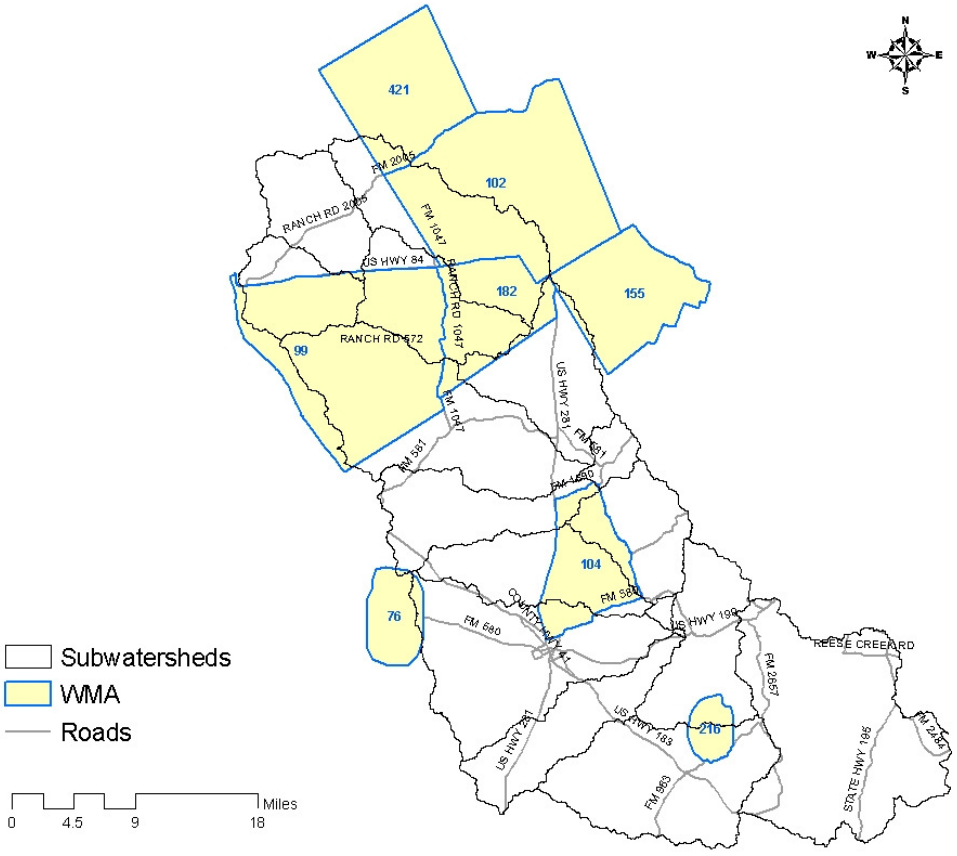

Figure 2.5. WMAs area locations in the Lampasas River watershed with deer population density estimations.

2.3.4.7. Feral Hogs

For feral hogs, a density of 13 hectares (32 acres) per animal unit was applied uniformly across forest, range, barren, crop, and managed pasture lands (Figure 2.4) within a 100 m buffer around the stream network of the watershed. An estimated total

population of 24,263 feral hogs was used with equation from Table 2.1 to estimate the daily potential *E. coli* loads resulting from feral hogs.

2.4. Results and Discussion

The spatial watershed analyses done using SELECT highlights subwatersheds that had the highest potential to contribute *E. coli* loads into a waterbody based on land use characteristics and pollutant contributor populations.

2.4.1. Spatially Explicit *E. coli* Load Estimation for the Buck Creek Watershed

Cattle are potentially the largest contributors of *E. coli* bacteria in the Buck Creek watershed while deer contribute the lowest *E. coli* load (Table 2.2). Best management practices should be applied for cattle and feral hogs since these are the largest potential contributors in the watershed.

Table 2.2. Source specific potential *E. coli* load ranges for the Buck Creek watershed.

Potential *E. coli* sources	**Range of Daily Potential *E. coli* Load (CFU/day)**
Cattle (Pasture and Range Cattle)	2.23×10^{12} to 4.20×10^{13}
Deer	1.69×10^{10} to 1.06×10^{11}
Feral Hogs	5.31×10^{11} to 4.10×10^{12}

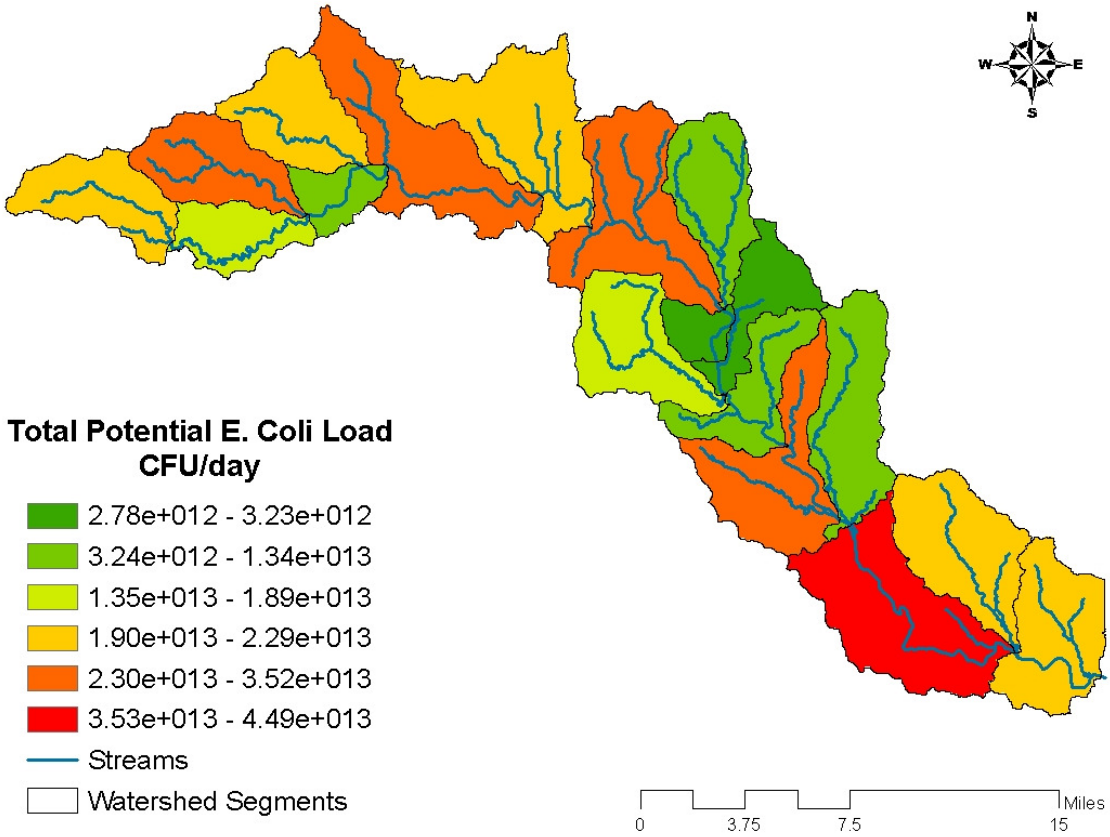

Figure 2.6. Total daily potential *E. coli* load resulting from various sources in the Buck Creek watershed.

Figure 2.6 illustrates the total potential load or the combined load which includes loading potentials from cattle, deer and feral hogs. Subwatersheds colored in red indicate areas with the highest potential for *E. coli* contributions to the creek while the darkest green areas represent areas with the lowest potential. The spatial analysis of *E. coli* sources shown in Figure 2.6 and the following figures are largely determined by the dominant land use in each subwatershed. For example, those areas dominated by cropland will have a lower potential for *E. coli* load than subwatersheds that are dominated by riparian forests or rangeland.

2.4.2. Spatial Distribution of *E. coli* Sources in the Little Brazos River Watershed

Table 2.3 illustrates that cattle are the highest contributors for all five of the Little Brazos tributary watersheds. Feral hogs are the second highest contributing potential source across all of the watersheds. Poultry operations are a higher contributor than feral hogs in the watersheds they are located in. On-site wastewater treatment systems are a significant contributor in the subwatersheds where there are hot spots for on-site wastewater treatment systems. Deer and wastewater treatment plants are the lowest contributing potential sources with wastewater treatment plants being the lowest.

Table 2.3. Source specific potential *E. coli* load ranges for the five tributaries of the Little Brazos River watershed.

Watershed	**Potential *E. coli* sources**	**Range of Daily Potential *E. coli* Load (CFU/day)**
Walnut Creek	Cattle	2.30×10^{9} to 3.36×10^{14}
	Deer	1.05×10^{6} to 8.97×10^{10}
	Feral Hogs	0 to 5.78×10^{12}
	Poultry Operations	1 to 6.37×10^{13}
	OWTSs	9.69×10^{6} to 5.41×10^{11}
	WWTFs	1 to 1.05×10^{9}
Mud Creek	Cattle	1.30×10^{14} to 2.55×10^{14}
	Deer	3.68×10^{10} to 7.37×10^{10}
	Feral Hogs	2.22×10^{12} to 3.98×10^{12}
	Poultry Operations	1 to 9.37×10^{12}
	OWTSs	6.15×10^{6} to 2.53×10^{12}
	WWTFs	1 to 1.43×10^{9}
Pin Oak Creek	Cattle	1.73×10^{13} to 1.09×10^{14}
	Deer	6.29×10^{9} to 3.33×10^{10}
	Feral Hogs	7.73×10^{11} to 2.08×10^{12}
	OWTSs	2.25×10^{10} to 4.63×10^{11}
Spring Creek	Cattle	3.58×10^{13} to 7.40×10^{13}
	Deer	1.37×10^{10} to 2.99×10^{10}
	Feral Hogs	9.70×10^{11} to 1.79×10^{12}
	OWTSs	6.07×10^{10} to 2.67×10^{11}
Campbells Creek	Cattle	4.80×10^{12} to 6.64×10^{13}
	Deer	1.81×10^{9} to 2.70×10^{10}
	Feral Hogs	1.31×10^{11} to 2.05×10^{12}
	OWTSs	4.25×10^{9} to 1.72×10^{12}

The subwatersheds of Walnut Creek Watershed had total potential *E. coli* loads between the "medium" and "high" ranges (Figure 2.7). This is mainly because of the size of the Walnut Creek Watershed and the amount of suitable areas of various contributing sources in comparison to the other watersheds in the Little Brazos Watershed. The subwatersheds of the Mud Creek Watershed had total potential *E. coli* load between "medium" and "high" range(Figure 2.7). With the size of the watershed being smaller, these results indicate Mud Creek as a high potential contributor of bacterial contamination to the Little Brazos River in comparison with these five watersheds. The subwatersheds of Pin Oak Creek Watershed had total potential *E. coli* loads between "low" and "medium" range (Figure 2.7). These results indicate Pin Oak Creek as a low potential contributor of bacterial contamination to the Little Brazos River in comparison with these five watersheds. All the subwatersheds of Spring Creek Watershed had total potential *E. coli* loads in "medium" range (Figure 2.7). These results indicate Spring Creek as a significant potential contributor of bacterial contamination to the Little Brazos River in comparison with these five watersheds. The subwatersheds of Campbells Creek Watershed had total potential *E. coli* loads between "very low" and "medium" range (Figure 2.7). These results indicate the bacterial contribution of Campbells Creek into the Little Brazos River being very low. However, the smaller size of the subwatersheds in Campbells Creek in comparison to the subwatersheds in the other five watersheds may skew the results somewhat because there is a lesser amount of area to be considered suitable for potential contributors

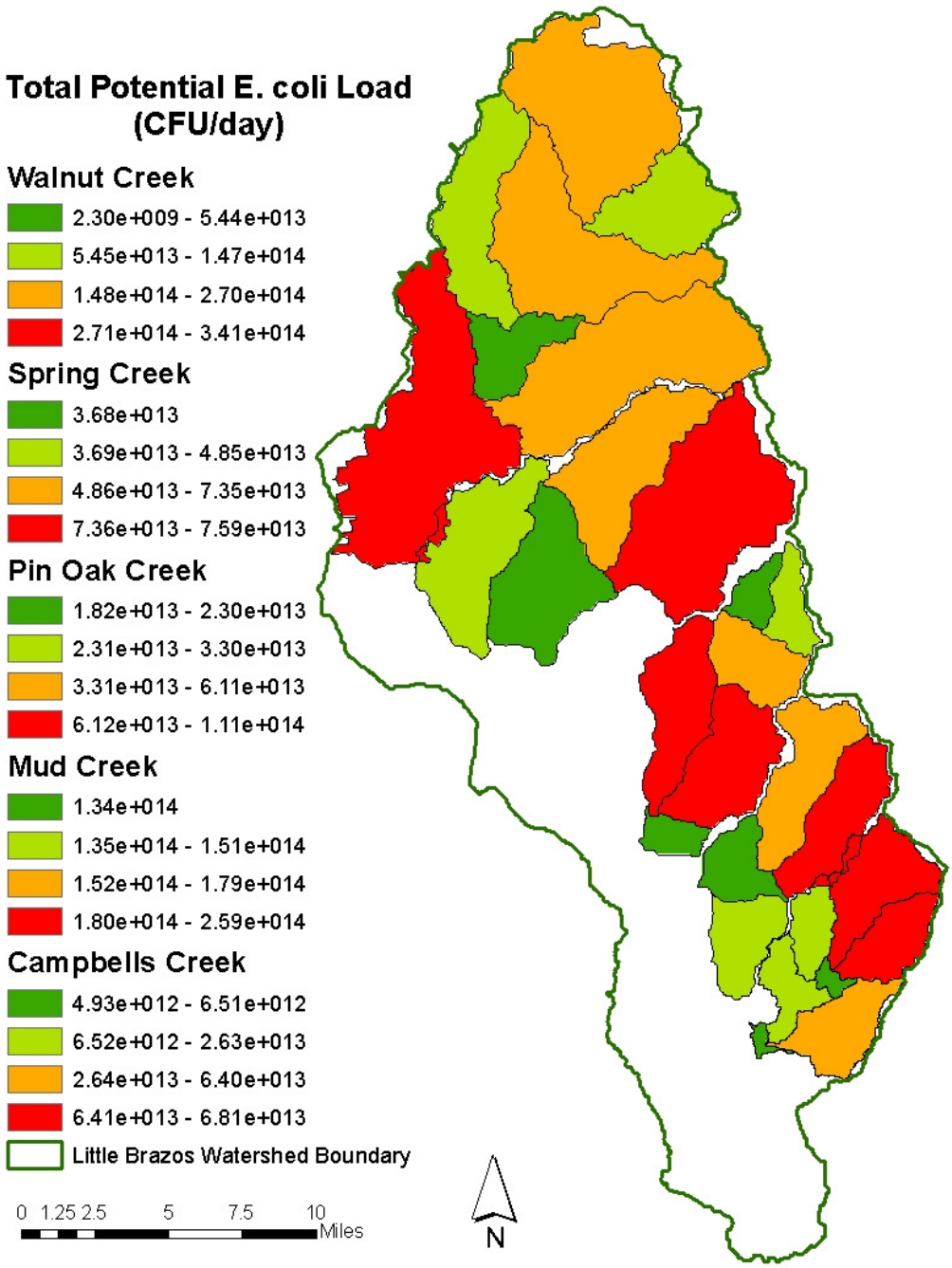

Figure 2.7. Total daily potential *E. coli* load resulting from various sources in the five tributary watersheds of the Little Brazos River watershed.

2.4.3. Total Daily Potential *E. coli* Loads Resulting from Various Sources in the Lampasas River Watershed as Predicted by SELECT

Table 2.4 illustrates the source specific *E. coli* ranges which can help to determine the amount each source is contributing to the watershed. The largest contributor for the Lampasas River watershed is cattle with feral hogs being the second largest contributor. OWTSs and dogs are also high contributors. CAFOs contribute more than feral hogs in the subwatersheds where they are present. Goats, sheep, and deer are not significant contributors and they contribute *E. coli* loads in close ranges. The least contributors are horses and WWTFs. Best management practices (BMPs) should be applied for cattle, feral hogs, and confined animal feeding operations because they are the largest contributors of *E. coli* loads in the watershed.

Table 2.4. Source specific potential *E. coli* load ranges for the Lampasas River watershed.

Potential *E. coli* sources	**Range of Daily Potential *E. coli* Load (CFU/day)**
Cattle	6.09×10^{13} to 3.91×10^{14}
Horses	8.36×10^{9} to 8.47×10^{10}
Goats	1.83×10^{12} to 9.56×10^{12}
Sheep	1.31×10^{12} to 8.18×10^{12}
Deer	1.04×10^{12} to 4.04×10^{12}
Feral Hogs	4.65×10^{12} to 1.86×10^{13}
On-site Wastewater Treatment Systems	3.24×10^{11} to 1.24×10^{13}
Wastewater Treatment Facilities	0 to 1.19×10^{10}
Dogs	2.25×10^{11} to 1.06×10^{13}
Confined Animal Feeding Operations	0 to 3.20 10^{13}

Figure 2.8 illustrates the Total Potential Load or the combined load which includes loading potentials from all of the contributing sources applied for the model. Subwatersheds colored in red indicate areas with the highest potential for *E. coli* contributions to the creek while the darkest green areas represent areas with the lowest potential. The subwatershed considered to be the highest contributor in the Lampasas River watershed as predicted by SELECT is most likely due to the large size of the subwatershed in comparison to the other subwatersheds.

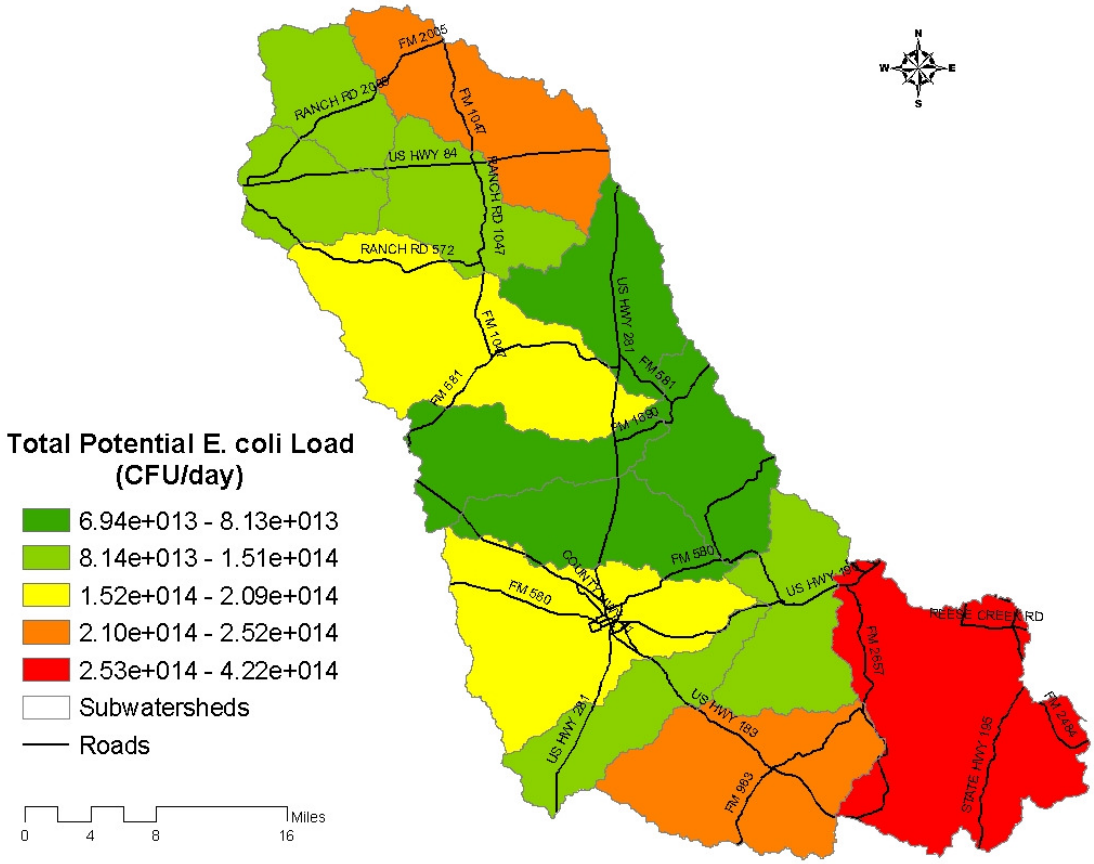

Figure 2.8. Total daily potential *E. coli* load resulting from various sources in Lampasas River watershed.

2.5. Conclusions

The SELECT methodology was applied to three rural watersheds: Buck Creek, Little Brazos River, and Lampasas River that are located in different regions of Texas. The SELECT methodology was able to be adapted for each watershed individually based on perceived potential contributing sources and data availability. Cattle were considered the highest contributor for all three watersheds and best management practices should be implemented to reduce the cattle contribution to the waterbodies. The SELECT methodology was able to not only highlight which contributing sources are of most concern but to also highlight the areas of highest concern to more effectively apply these best management practices. The SELECT methodology can be easily adapted and applied to watersheds to reflect stakeholder knowledge and concerns.

CHAPTER III

ESTIMATING *E.COLI* CONCENTRATIONS FROM NON POINT SOURCES FOR GERONIMO CREEK

3.1. Introduction

When developing a Watershed Protection Plan (WPP) or a Total Maximum Daily Load (TMDL), it is often difficult to accurately assess the pollutant load for a watershed as a result of inadequate water quality monitoring data. Bacteria are the most common reason for impairment of Texas water bodies. According to the Texas Commission on Environmental Quality (TCEQ), there are 274 bacteria impairments in Texas water bodies out of 386 impaired water bodies (TCEQ, 2008). Data on bacteria in water bodies is often more sparse than other types of water quality data, which hinders the development of WPPs or TMDLs.

In order to develop WPPs or TMDLs, additional data on waterborne bacteria must be collected which is costly and time consuming. The bacteria load analysis for a watershed cannot begin until the water quality monitoring data collection is completed. Generally, water quality data can take anywhere from a year to multiple years to collect for a substantial dataset. The U.S. EPA estimates water quality monitoring of all TMDLs nationally, "The cost of water quality monitoring to support the development of TMDLs is expected to be approximately $17 million per year" (USEPA, 2001b). A considerable portion of developing a TMDL is to allocate pollutant loads and to identify

potential sources. This can be done with modeling which can be costly and require a significant amount of input data.

Models such as Soil and Water Assessment Tool (SWAT) and Hydrological Simulation Program- FORTRAN (HSPF) have been used for bacterial modeling (Benham, et al., 2006; Sadeghi & Arnold, 2002). Other simplistic microbial models such as, the potential non-point pollution index (PNPI) and a Spatially Explicit Delivery MODel (SEDMOD) have been developed to rank the potential pollution impacts of areas from nonpoint sources primarily utilizing land use and geomorphology (Fraser, et al., 1998; Munafo, et al., 2005).

SELECT is an automated Geographic Information System (GIS) tool that can assess potential *E. coli* loads in a watershed based on spatial factors such as land use, population density, and soil type (Teague, et al., 2009). SELECT is able to calculate a potential *E. coli* load and highlight areas of concern for best management practices (BMPs) to be implemented. The potential *E. coli* load in SELECT is calculated by distributing the contributing sources spatially over the entire watershed. The population densities of potential contributors are determined with stakeholder input to accurately represent the watershed; however, SELECT is a worst case scenario model and assumes the largest amount of contribution possible from individual sources.

Current bacteria models either require extensive monitoring data within the watershed for calibration or are not able to predict actual *E. coli* concentrations in the waterbody. A simple model that is able to predict actual bacteria concentrations in a waterbody is needed in order to develop TMDLs or WWPs within the State of Texas.

The objective of this study was to develop a model that would estimate the runoff volume and the *E. coli* concentration contributed by surface runoff at a sampling site drainage area outlet.

3.2. Study Area

Located in the Guadalupe River basin, the Geronimo Creek watershed is located across Comal and Guadalupe Counties in south central Texas (Figure 3.1). The Geronimo Creek watershed consists of Geronimo Creek and its tributary, Alligator Creek. Alligator Creek is an intermittent stream that typically only has flow after a rainfall event. Geronimo Creek is a tributary of the Guadalupe River which is used for recreation by local residents and tourists. The watershed is 17,868 hectares (44,152 acres) and is primarily agricultural with some urban near the towns of Seguin and New Braunfels (Dictson, 2009).

Geronimo Creek was chosen as the study site because it is listed as a bacterially impaired waterbody on the 303(d) list (TCEQ, 2008). A WPP for Geronimo Creek is also being developed by the Texas AgriLife Extension Service – Department of Soil and Crop Sciences through a Texas State Soil and Water Conservation Board project with the Clean Water Act 319(h) Non Point Source Grant Program. SELECT modeling was performed to assess the potential *E. coli* loads to develop the load allocation portion of the WPP.

The WPP project enables us to receive crucial data such as potential sources, population densities of animals, and what areas or land use types where the sources would be present from a local stakeholder group consisting of affected and owners and citizens.

3.2.1. Samples

Historical and routine stream flow and *E. coli* concentration sampling data ranging from 1996 to 2010 were obtained from the Guadalupe Brazos River Authority (GBRA). The SH 123 and Haberle Road sampling sites are both historical sites while the other 13 samplings sites in the watershed began sampling in September 2008. Haberle Road samples were taken on a monthly basis beginning in September 2003 and ending in December 2010 resulting in 84 samples. For the SH 123 sampling site, monthly sampling began in October 1996 and ended in August 2003, but then resumed on September 2008 until August 2010. Out of the 105 data points taken at the SH 123 sampling site only 5 coincided with precipitation induced runoff. Only 12 data points out of the 84 for the Haberle Road site samples were taken when precipitation induced runoff occurred.

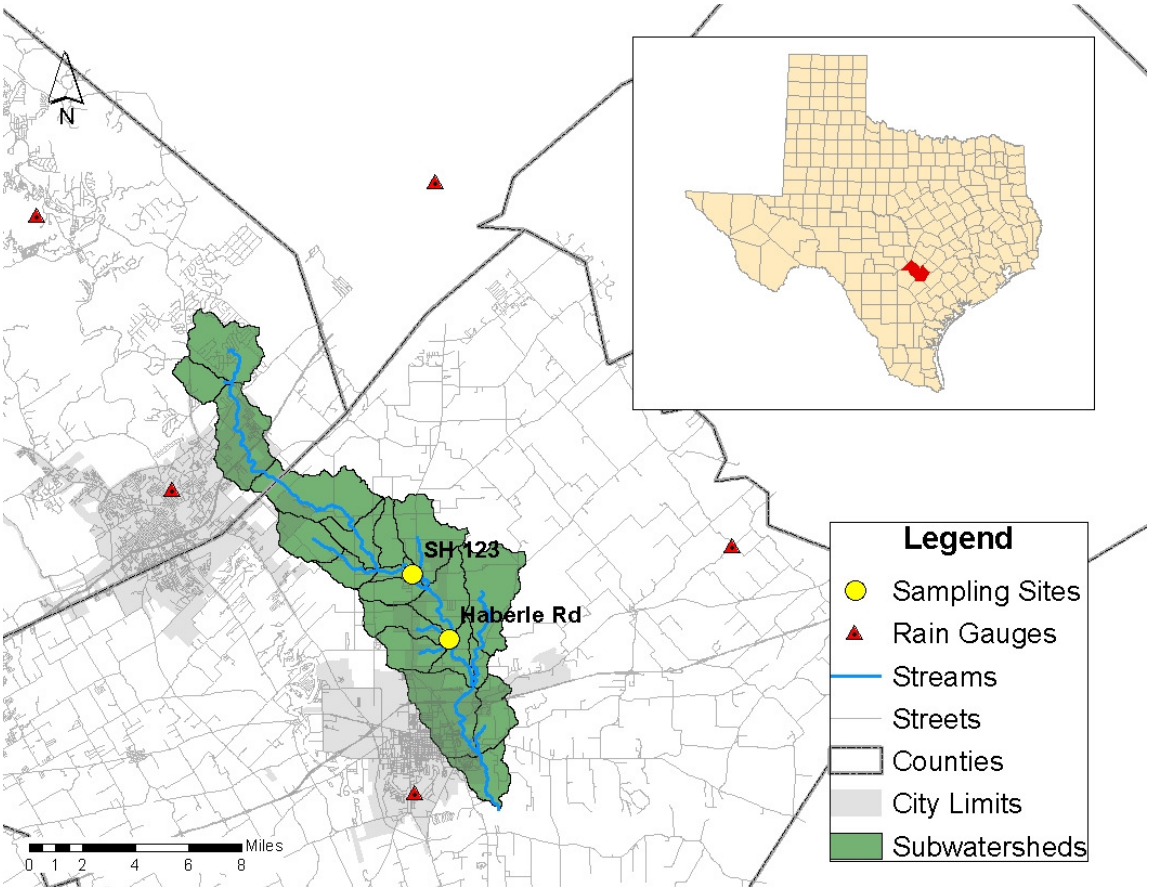

Figure 3.1. Geronimo Creek watershed study area with NCDC rain gauges and water quality sampling sites.

3.3. Methodology

E. coli concentrations were calculated using a modified delivery factor originally developed by McElroy et al. (1976) for pollutant loading from livestock facilities:

$$C = (Y * D)/(a * R * A) \tag{3.1}$$

where,

C = concentration of *E. coli* at sampling site (CFU/mL)

Y = daily loading rate of *E. coli* at sampling site (CFU)

D = delivery factor (dimensionless)

a = unit conversion factor (2.54×10^4) – to convert from in•m^2 to mL

R = daily runoff at sampling site (in)

A = grid cell area (m^2)

Equation 3.1 was intended for livestock facilities but was applied to multiple non-point sources for this research calculated using SELECT and ArcGIS 9.X (McElroy et al., 1976). The variable concentration of pollutant in runoff (C) was calculated using the equation above to determine the concentration of *E. coli* in Geronimo Creek. The loading rate (Y) was calculated in SELECT for livestock, wildlife, and domestic sources. McElroy et al. (1976) acknowledged that the quantity of pollutants discharged depends mostly on runoff volume. The runoff volume (R) was calculated with an automated program developed in ArcGIS 9.3. using the SCS curve number approach with daily precipitation data.

3.3.1. Runoff (R)

Daily precipitation data was obtained at 5 sites, Canyon Dam, Kingsbury, New Braunfels, San Marcos, and Seguin, from the National Climatic Data Center (NCDC) for 1996 to 2010. The NCDC rain gauges shown in Figure 3.1 were utilized to develop a daily precipitation grid using inverse distance weighted (IDW) interpolation over the entire watershed area with a grid cell size of 30 meters. The cell size of 30 meters was used to maintain a constant cell size with the SELECT results. There are rain gauges located in the watershed from the Guadalupe-Blanco River Authority (GBRA) but these were not used because the data was not consistent and was only available from 2004 to 2008.

The minimum rainfall needed to produce runoff was calculated using the SCS curve number approach by first calculating the equation for the maximum soil water retention parameter (*S*):

$$S = 1000/CN - 10 \quad (3.2)$$

where,

S = maximum soil water retention parameter (in)

CN = area weighted curve number for the Geronimo Creek watershed

S multiplied by 0.2 is the minimum amount of rainfall required in the watershed to produce runoff. This rainfall amount was used to determine the days where runoff precipitation and a sampling event occurred simultaneously in the watershed. These days were the days where the model was run to prevent running the model on days where no runoff was generated. A custom land use classification (Figure 3.2) was provided by the Texas A&M University Spatial Sciences Laboratory (SSL) using 2008 National Agriculture Imagery Program (NAIP) imagery and a prior Texas Parks and Wildlife (TPWD) Classification. The watershed curve number grid (Figure 3.3) was developed in ArcGIS 9.X. by intersecting the Soil Survey Geographic (SSURGO) hydrologic soil group with the land use type and using an NRCS Curve Number Lookup Table (Soil Conservation Service, 1986).

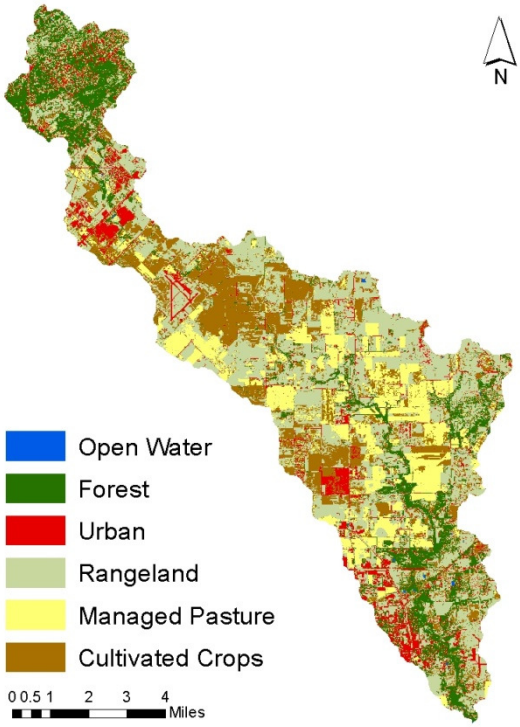

Figure 3.2. Geronimo Creek watershed land use classification.

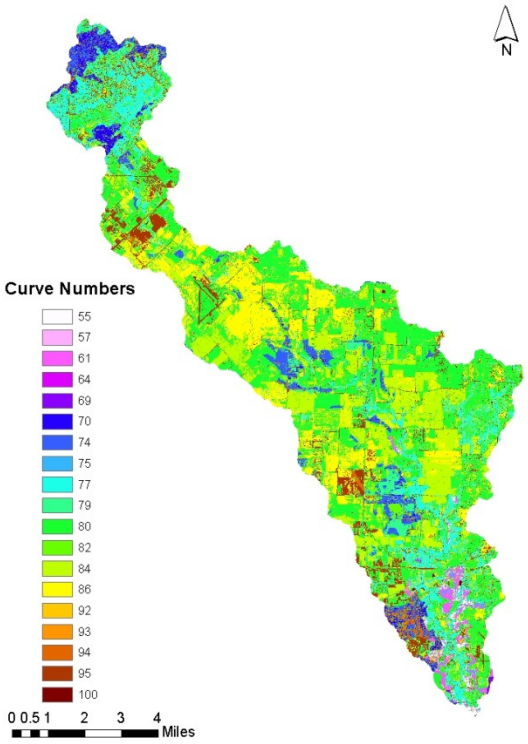

Figure 3.3. Curve number grid.

Table 3.1. NRCS curve number lookup table.

Land Use Type	Hydrologic Soil Group	Curve Number
Open Water	A	100
	B	100
	C	100
	D	100
Forest	A	25
	B	55
	C	70
	D	77
Urban	A	89
	B	92
	C	94
	D	95
Rangeland	A	39
	B	61
	C	74
	D	80
Managed Pasture	A	49
	B	69
	C	79
	D	84
Cultivated Crops	A	65
	B	75
	C	82
	D	86

The curve numbers (Table 3.1, Figure 3.3) used in the NRCS lookup table (Soil Conservation Service, 1986) were determined based on the assumption of a normal antecedent moisture condition, i.e AMC II.

The area weighted curve number for the Geronimo Creek Watershed was calculated as 82. The minimum rainfall needed to produce runoff calculated using the area weighted curve number was 0.44 inches.

If one of the five rain gauges measured precipitation greater than the minimum rainfall to induce runoff, a precipitation grid was developed using the ArcGIS Spatial Analysis Extension on each day that routine *E. coli* samples were taken at the Geronimo Creek sampling sites. Interpolation was done by the IDW technique for rainfall depths across the watershed; inverse distance weighting assumes that observations closer to one another are more alike than ones farther apart (Zhang & Srinivasan, 2009).

The runoff volume at a sampling site was calculated from the precipitation grid (Figure 3.4). An automated tool was programmed into ArcGIS to calculate a runoff grid with the input of a rain gauge shapefile with the measured amounts of rainfall for each rain gauge as fields in the attribute table and an S grid calculated from the curve number grid.

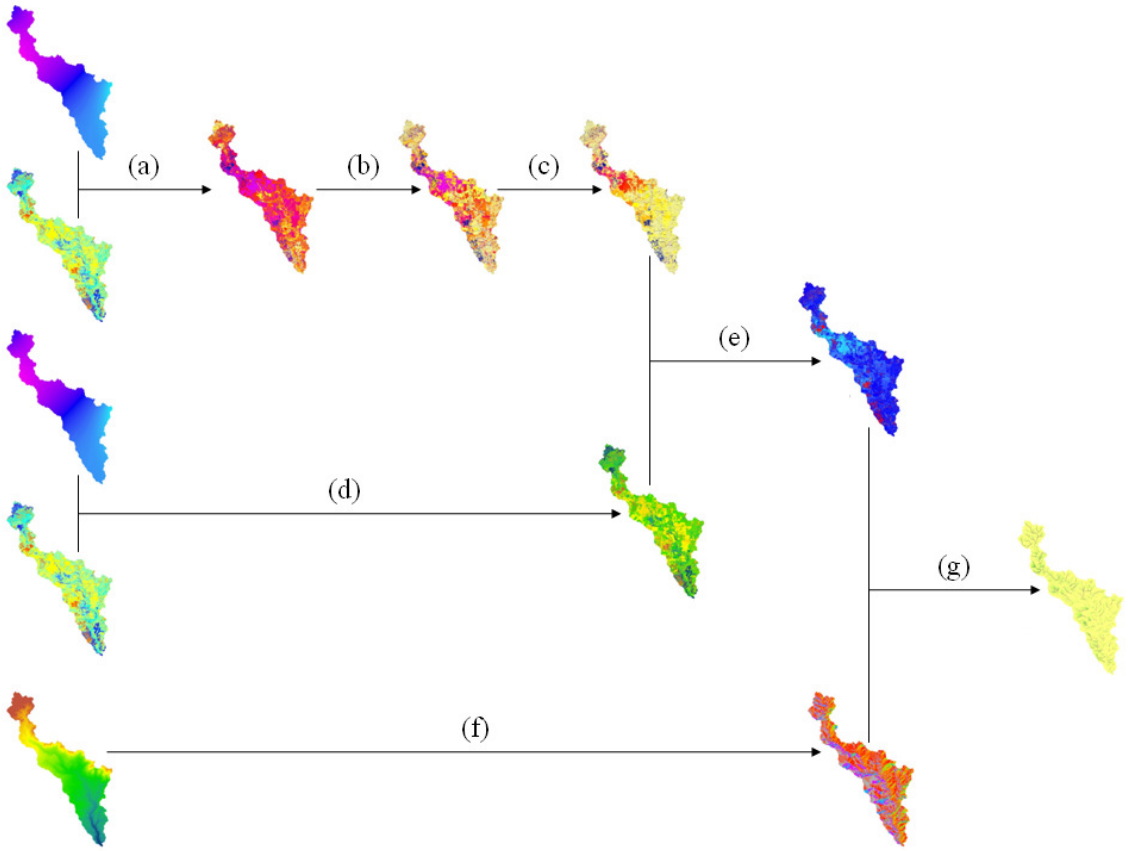

Figure 3.4. Flow chart illustrating the calculation of accumulated runoff volume. (a) Subtract 0.2S from rainfall grid, (b) Convert negative values to zero, (c) Square corrected grid, (d) Add 0.8S to rainfall grid (e) Divide grid created from (c) by grid created from (d) then multiply result by 900 square meters to create a runoff volume grid, (f) Compute flow direction from DEM grid, (g) Compute flow accumulation from flow direction grid using the runoff volume grid as the accumulation weight.

The runoff volume grid (Figure 3.4(a-e)) was calculated using the SCS curve number approach with the equation:

$$Q = [(P - 0.2S)^2/(P + 0.8S)] * A \quad (3.3)$$

where,

Q = runoff volume (in-m^2)

P = precipitation (in)

S = the maximum soil water retention parameter (in)

A = area of a grid cell (m^2)

The curve number grid is calculated into an S grid using equation 3.2. Equation 3.3 requires that P must exceed 0.2S before any runoff is generated; therefore, when cells in Figure 3.4(a) resulted in negative values, they are given a value of zero (Figure 3.4 (b)) so that runoff was not calculated for cells with P less than 0.2S. The result of Figure 3.4 (b) was then squared creating the numerator of Equation 3.3(Figure 3.4 (c)). The denominator of equation 3.4 was calculated by adding 0.8S to P (Figure 3.4(d)). The numerator (Figure 3.4(c)) was then divided by the denominator (Figure 3.4(d)) which calculated the runoff depth in inches for every grid cell in the entire watershed. Runoff depth was then converted to a runoff volume per grid cell by multiplying by the cell area, 900 m^2 (Figure 3.4(e)).

An additional part of the Arc GIS 9.3. tool was used to automatically calculate the runoff volume accumulation grid for the watershed shown in Figure 3.4(f-g). The inputs to the tool were the previously generated runoff volume grid and a Digital Elevation Model (DEM) over the watershed area which had a 30 X 30 meter grid cell

size provided by the Texas A&M University SSL. The result of the flow accumulation is the total amount of runoff volume going through a specific grid cell. The runoff volume at each sampling site was estimated by identifying the runoff volume value at each sampling site drainage area outlet.

3.3.2. Potential *E. coli* Load (*Y*) Estimation Using SELECT

Potential *E. coli* loads for Geronimo Creek were predicted using SELECT and input from stakeholders for stocking rates and possible sources. A custom land use classification (Figure 3.4) provided by the Texas A&M University SSL was used to distribute animals on land use types that were determined to be suitable for a specific animal or source.

Twenty-one subwatersheds were delineated using the SWAT model. The stream channel was determined with the SWAT model as well using the DEM. In the Geronimo Creek watershed, it was determined that livestock sources for the watershed are goats, horses, and cattle. Wildlife sources are deer and feral hogs. Domestic sources consist of dogs and on-site wastewater treatment systems (OWTSs). The fecal production rates used in the model from the EPA guidance are in fecal coliform (USEPA, 2001a). These rates then need to be converted from fecal coliform to *E. coli*. A conversion of 0.63 fecal coliform to *E. coli* was used in the model. The conversion factor of 0.63 was decided using the USEPA's regulatory standards for fecal coliform and *E. coli* in recreational waters. The regulatory standard for fecal coliform was 200 organisms per 100 mL and is 126 organisms per 100 mL for *E. coli* (USEPA, 2003).

The conversion factor was determined by taking the ratio of these two regulatory standards.

For livestock and wildlife, the number of animals is estimated with animal densities and stakeholder input. For cattle, the stakeholders determined stocking rates of 8 and 4 hectares (20 and 10 acres) per animal should be applied to Comal and Guadalupe Counties, respectively, and should be applied to the land use types of rangeland, forest, and managed pasture which were determined as suitable habitats for cattle in this area. A density for horses was determined to be 53 hectares (132 acres) per animal over the entire watershed with a total watershed population of 124 horses with a suitable habitat of rangeland. The animals are distributed evenly across suitable habitats and a fecal production rate is then applied per animal. Since goats are typically raised on goat farms, stakeholders determined that 200 goats out of the total watershed population of 750 animals should be distributed evenly in the watershed on rangeland, forest, and manage pasture land use types. The remaining animals were concentrated to specific watersheds which contained known goat farms for a specified number of animals. The potential *E. coli* load for the subwatersheds containing goats was calculated per subwatershed by multiplying the number of animals per subwatershed by the fecal production rate per animal. White-tailed deer had a population density of 4 hectares (10 acres) per animal (Lockwood, 2005). The suitable habitat determined for deer were forest and rangeland with at least 8 hectares (20 acres) of contiguous terrain available. Feral hogs had a population density of 11 hectares (26 acres) per animal and were only distributed on suitable habitat within 100 meters of the main stem of Geronimo Creek

which is perennial. Feral hogs were not distributed around Alligator Creek because it is an intermittent creek and is an unsuitable habitat for feral hogs. The suitable habitats for feral hogs as determined by stakeholders were forest, rangeland, managed pasture, and cultivated crops.

For dogs, the 2000 census data was used to calculate the contribution by using a dog density of 1 dog per household. The potential *E. coli* load for OWTSs was calculated by Espey Consultants. For OWTSs, spatially distributed point data of each household was collected from 911 address data and households within Certificate of Convenience and Necessity (CCN) areas were removed to not include households being serviced by a wastewater treatment facility. CCN areas are on city sewer lines and therefore, the waste is treated at a WWTF and not by OWTSs. A failure rate was determined for the OWTSs using SSURGO soil limitation classes and the age of the system to calculate the percentage of *E. coli* contributing to the watershed due to septic failure. A fecal production rate was then applied to each household for dogs and OWTSs. Since SELECT divides the watershed into a raster grid with a 30 X 30 meter cell size, the potential load is calculated over the entire watershed at a 30 X 30 meter cell size. The individual raster files for each source are then added together spatially to create a total load raster (Figure 3.5) for the watershed that is divided into 30 X 30 meter grid cells.

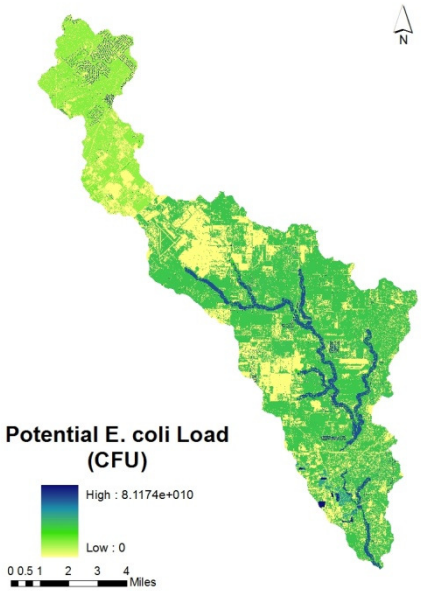

Figure 3.5. Total potential *E. coli* load calculated using SELECT for the Geronimo Creek watershed.

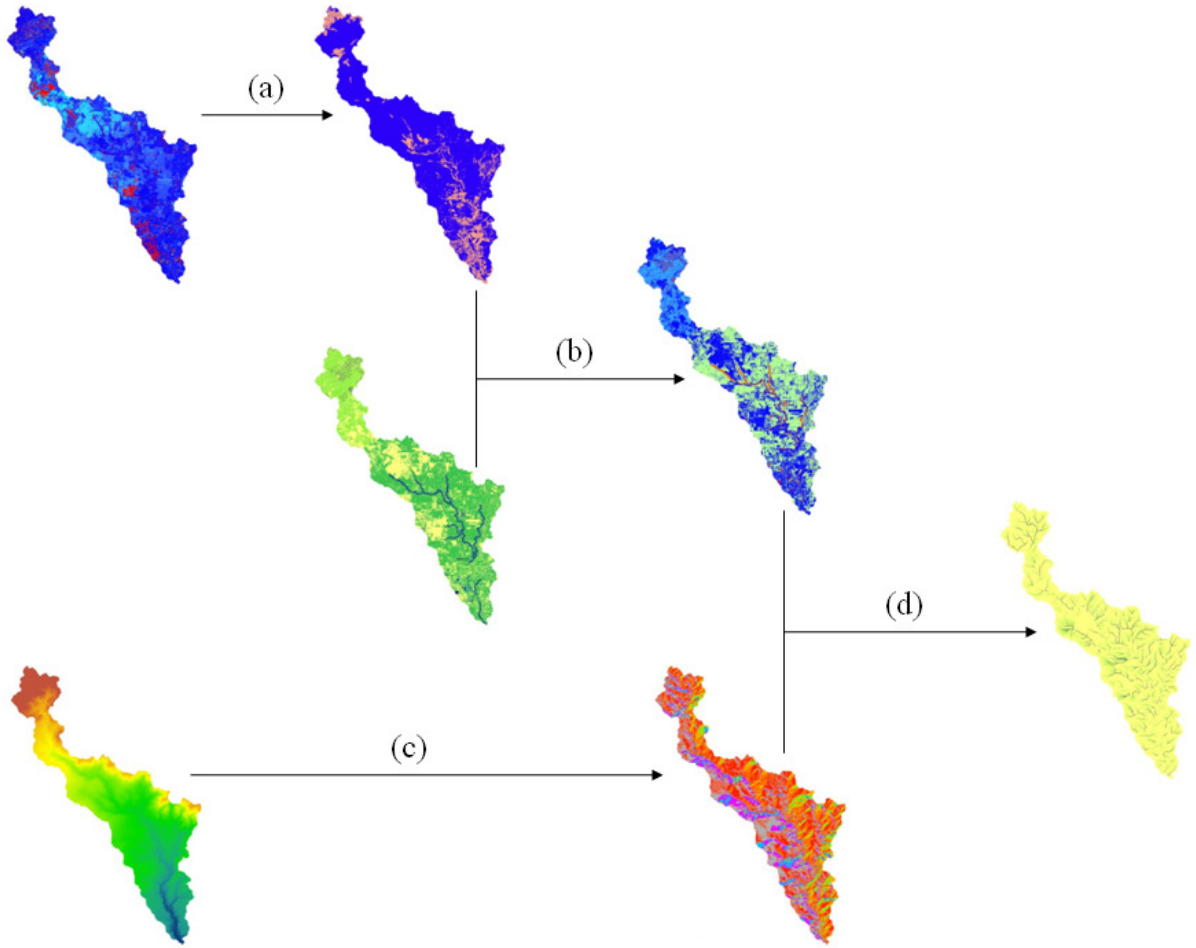

Figure 3.6. Flow chart illustrating the calculation of the contributing *E. coli* load. (a) Convert runoff values greater than zero to one (b) Multiply converted runoff and SELECT load to compute contributing load (c) Compute flow direction (d) Compute flow accumulation using flow direction with contributing load as accumulation weight.

The total load raster (Figure 3.5) estimates the potential *E. coli* load for the entire watershed based on a worst case scenario assuming the entire load calculated reaches the water body. Another part of the tool programmed in ArcGIS 9.X. was to calculate the *E. coli* load actually reaching a specific grid cell in the watershed (Figure 3.6).

The inputs to the tool were the previously calculated runoff grid, the total load raster from SELECT, and the DEM. The first step shown in Figure 3.6 to estimate the *E. coli* load reaching the sampling site was to only consider the *E. coli* load grid cells that have runoff generated. The runoff volume grid generated from each precipitation event was converted to a Boolean runoff grid, where the grid cells with values greater than 0 were converted to 1. A runoff SELECT grid (Figure 3.6(b)) was estimated for each runoff event by multiplying the SELECT grid by the Boolean runoff grid, so that the cells with no runoff generated had a contributing *E. coli* load of zero. The load accumulation was calculated using the runoff SELECT grid as an input weight and the DEM shown in Figure 3.6 (c-d). The output of flow accumulation would then represent the amount of *E. coli* load that would flow through each cell considering the upslope cells. The flow accumulation at a sampling site would then estimate the *E. coli* load reaching that site.

3.3.3. Calculation of Observed and Predicted Runoff Volumes

The observed instantaneous stream flows corresponding to the time each *E. coli* sample was taken were converted to runoff volumes. Two flow duration curves (FDCs) were developed for each site, Haberle Road and SH 123, one from instantaneous monthly samples (Figure 3.7 and Figure 3.8) and the other from SWAT simulated daily

flow rates (Figure 3.9 and Figure 3.10). For SH 123 and Haberle Road, the FDCs were developed using daily SWAT simulated flowrates from 1998 to 2009. The FDCs developed for SH 123 and Haberle Road using instantaneous flowrates from 1996 to 2003 with a break in sampling until 2008 to 2009 and from 2003 to 2009, respectively.

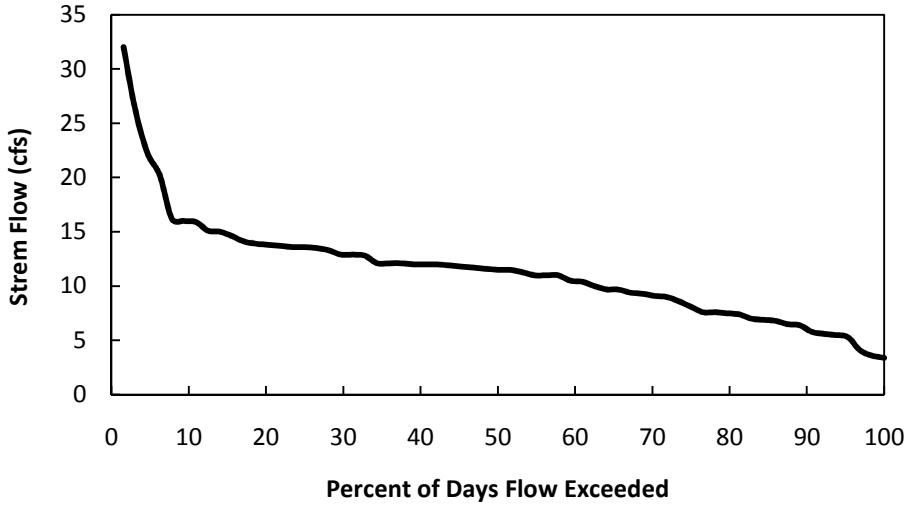

Figure 3.7. Haberle Road FDC developed using instantaneous flows.

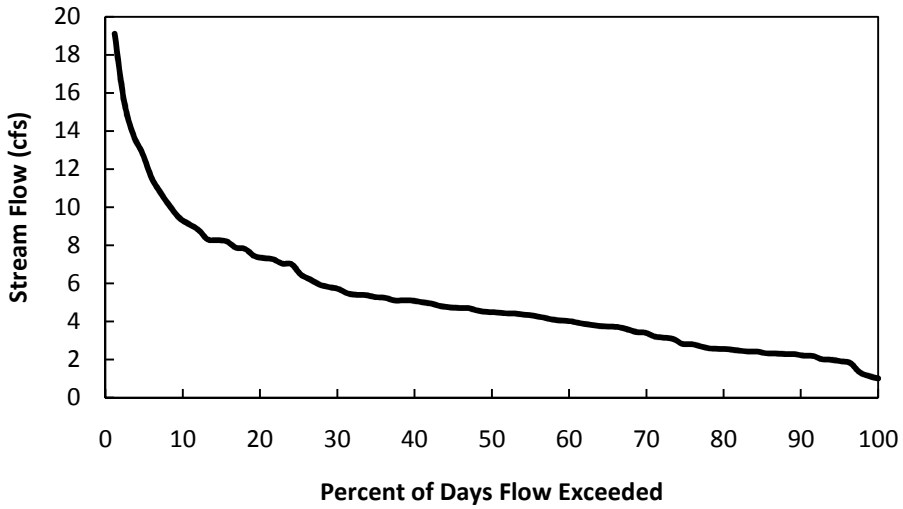

Figure 3.8. SH 123 FDC developed from instantaneous flows.

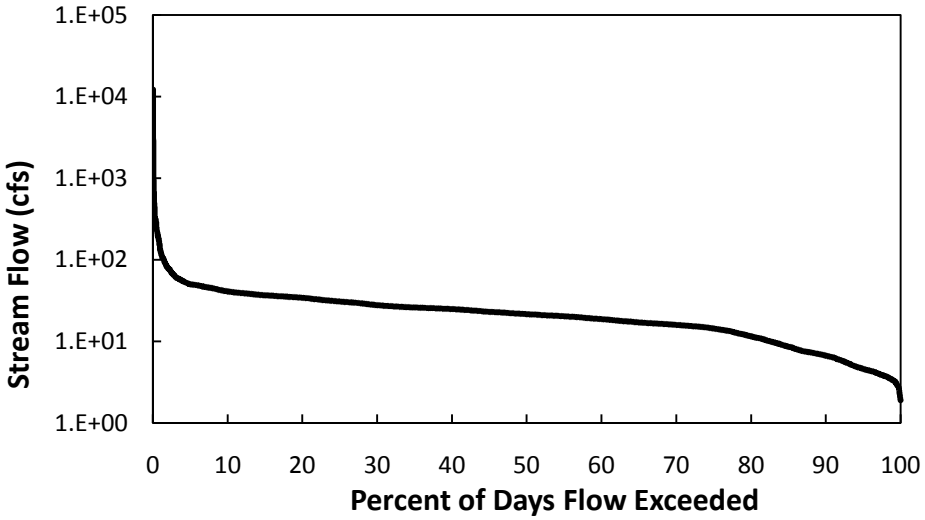

Figure 3.9. Haberle Road FDC developed from SWAT simulated flows.

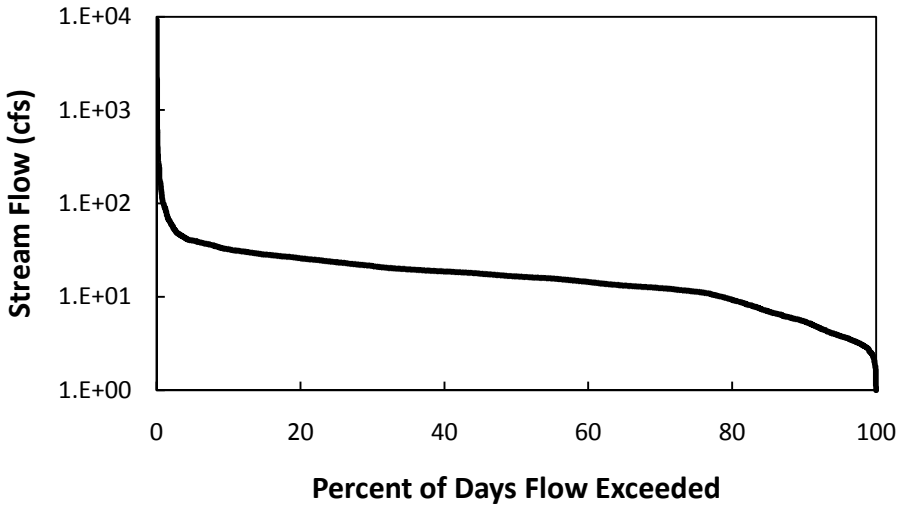

Figure 3.10. SH 123 FDC developed from SWAT simulated flows.

Three methods were used for each site to add base flow to the predicted runoff volume: adding the 75% exceedence flow calculated from the FDC developed using instantaneous flows, the 75% exceedence flow calculated from the FDC developed using SWAT simulated flows, and the 100% exceedence flow developed using instantaneous flows. Flow duration curves were developed for the sampling sites SH 123 and Haberle Road using SWAT simulated flow rates ranging from 1998 to 2009. The 75% exceedence flow using instantaneous flows were 2.2 cfs for SH 123 and 8.1 cfs for Haberle Road. The FDCs developed using SWAT simulated flows for the 75% exceedence flow were 11.35 cfs and 14.29 cfs for SH 123 and Haberle Road, respectively. The 100% exceedence flow for SH 123 and Haberle Road were 1.0 cfs and 3.4 cfs, respectively. All of the exceedence flows were converted to volumes using the same method and time as the observed flows.

The stream flow (cfs) was converted to a runoff volume (m^3) by multiplying by time and conversion factors. The times were calculated using multiple methods shown in Table 3.2: SWAT calculated time of concentration, SWAT calculated lag time, manually calculated lag time, and manually calculated time of concentration. These times were used to more accurately determine the amount of water flowing through a sampling site from a rainfall event rather than assuming 24 hours. The SWAT calculated time of concentration was calculated for each hydrological response unit (HRU). All of the HRUs in a subbasin were then averaged together and the average time of concentration calculated for the subbasins containing the main stream channels of Geronimo and Alligator Creek were added together from the upper portion of the watershed until SH 123 or Haberle Road. The SWAT calculated lag time was converted from a time of concentration by using a method for lag time estimation developed by the SCS (Haan et al., 1994):

$$t_L = 0.6t_C \tag{3.4}$$

where,

t_L = lag time (hrs)

t_c = time of concentration (hrs)

Table 3.2. Various times used to estimate runoff volumes in the Geronimo Creek watershed.

Sampling Site	Method	Time (hrs)
SH 123	Lag Time Manually Calculated from Alligator Creek	7.2
	Time of Concentration Manually Calculated from Alligator Creek	12.0
	Time of Concentration SWAT Calculated from Alligator Creek	6.9
Haberle Road	Lag Time Manually Calculated from Alligator Creek	9.2
	Time of Concentration SWAT Calculated from Alligator Creek	7.8
	Lag Time Manually Calculated from SH 123	2.9
	Time of Concentration SWAT Calculated from SH 123	0.97
	Lag Time SWAT Calculated from SH 123	0.58

The manually calculated lag time used the SCS lag time equation based on natural watersheds (Haan, 1994):

$$t_L = L^{0.8}(S+1)^{0.7}/1900Y^{0.5} \quad (3.5)$$

where,

L = hydraulic length of the sampling site drainage area (ft)

S = average maximum soil water retention parameter (in)

Y = average land slope of the sampling site drainage area (%)

The SH 123 hydraulic length (L) was determined by measuring the longest distance along the SWAT delineated stream channel to the drainage area outlet. The stream length included the entire length of Alligator Creek and the length of Geronimo Creek from its confluence with Alligator Creek to the drainage area outlet. The manually calculated time of concentration was determined using equation 3.4. Although the Haberle Road sampling site is located downstream of SH 123, there is a log jam located at SH 123 which may be inhibiting flow from upstream of SH 123 to Haberle

Road. Times of concentration and lag times were calculated for Haberle Road for the entire upstream portion of the watershed and also only from SH 123.

3.3.4. Delivery Factor (*D*)

The delivery factor is back calculated from equation 3.1 using observed Guadalupe Blanco River Authority (GBRA) *E. coli* concentration data taken from each of the SH 123 and Haberle Road sampling sites. The delivery factor represents all factors influencing movement of the potential *E. coli* load into the creek with the exception of surface runoff. Two separate delivery factors were calculated, one using observed runoff volume, and the other calculated from the model predicted simulated runoff volume.

A delivery factor was calculated for each measured *E. coli* sample using both the observed and simulated runoff volumes for each site. For each site, the individual delivery factor values were averaged using both a mathematical mean and a geometric mean. This resulted in the calculation of eight different delivery ratios (Table 3.3) to be applied to the data.

Table 3.3. Different delivery factors used for *E. coli* concentration calculation.

Sampling Site	Average	Observed Runoff Volume	Predicted Runoff Volume
SH 123	Mean	0.752	0.942
	Geomean	0.015	0.015
Haberle Road	Mean	0.480	0.316
	Geomean	0.065	0.059

3.3.5. Statistics

The accuracy of the model was evaluated using the Nash-Sutcliffe efficiency (*E*), root mean square error (*RMSE*), and RMSE-observations standard deviation ratio (*RSR*). According to Nash and Sutcliffe (1970) the *E* value is an index of agreement or disagreement between observed and predicted values. The *E* value evaluates how consistently the predicted values agree with the observed values by applying linear regression analysis (Nash and Sutcliffe, 1970). *E* is computed with equation 3.6 (Nash and Suttcliffe, 1970):

$$E = 1 - [\sum_{i=1}^{n}(O_i - P_i)^2 / \sum_{i=1}^{n}(O_i - \bar{O})^2] \quad (3.6)$$

where,

O_i = observed values

P_i = predicted values

$\bar{O}$ = mean of the observed values

n = number of samples

The *E* value ranges from negative infinity to 1, where negative values are considered a biased model and values between 0 and +1 are considered an unbiased model (McCuen, et al., 2006). Model efficiencies were classificed similar to Moriasi et al. (2007) and Parajuli et al. (2009) as very good (E = 0.75 to 1), good (E = 0.5 to 0.74), fair (E = 0.25 to 0.49), poor (E = 0 to 0.24) and unsatisfactory ($E < 0.0$).

RMSE is an error index used in model evaluation and is valuable because the error is indicated in the units of the constituent of interest (Moriasi, et al., 2007). Legates and McCabe (1999) recommend including at least one relative error measure (*E*

or R^2) and at lease one absolute error measure (*RMSE* or mean absolute error) for a complete assessment of model performance. *RSME* values close to 0 indicate a perfect fit but values half the standard deviation are still considered low (Singh, et al., 2004). The equation for *RMSE* is:

$$RMSE = \sqrt{\sum_{i=1}^{n}(O_i - P_i)^2/n} \tag{3.7}$$

RSR is a model evaluation statistic that standardizes RMSE (Equation 3.7) with the observed data standard deviation (Moriasi, et al., 2007). Moriasi et al. (2007) developed *RSR* to fill the need of an error index with additional information provided for using *RSME* with the standard deviation recommended by Legates and McCabe (1999). The *RSR* is computed using equation 3.8 (Moriasi, et al., 2007):

$$RSR = \left[\sqrt{\sum_{i=1}^{n}(O_i - P_i)^2}\right] / \left[\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2}\right] \tag{3.8}$$

The value of *RSR* ranges from 0, which is the optimal value and indicates a perfect model, to a large positive value (Moriasi, et al., 2007). Model efficiences are classified by Moriasi et al. (2007) as very good (RSR = 0.00 to 0.50), good (RSR = 0.51 to 0.60), satisfactory (0.61 to 0.70), and unsatifactory (RSR > 0.70).

Moriasi et al. (2007) states that the model evaluation guidelines for both *E* and *RSR* values given apply to a continuous, long-term simulation for a monthly time step. The guidelines should be adjusted based on a multitude of factors including quality and quantity of measured data, single-event simulation, evalution time step, model calibration procedure, and project scope and magnitude (Moriasi, et al., 2007). Moriasi et al. (2007) continues to say that when a complete measured time series does not exist,

such as when only a few grab samples per year are available, that the data may not be sufficient for analysis using the recomended statistics.

3.4. Results and Discussion

The runoff volumes were simulated for both the SH 123 and Haberle Road sampling sites and compared to the instantaneous observed flows converted to a volume. For both SH 123 and Haberle Road, *E. coli* concentrations were simulated using equation 3.1 with both predicted simulated and instantaneous observed runoff volumes.

3.4.1. Outlier Testing

The model assumes that the runoff transporting the *E. coli* to the stream is generated by the rainfall occurring on the same day the sample is taken. An ideal sample would be a sample where either there was no rainfall or there was not enough rainfall to induce runoff from any site the day before the sample was taken. Eleven out of the 17 sites fit these guidelines and were considered ideal samples. Five of the other samples had runoff generated from one or more sites the day before the sample was taken, but the cumulative rainfall of all sites was less for the day before than the day the actual sample was taken. For the two cases where runoff was generated from multiple sites, the cumulative rainfall for the day the sample was taken was at least 4.5 times greater than the cumulative rainfall of all sites the day before the sample was taken. It is assumed that these 16 samples would therefore, not impact the model significantly because a majority of the rainfall impacting runoff is occurring on the same day the sample is taken. One sampling day, 3/3/05, for the Haberle Road site did not follow these guidelines. The cumulative rainfall of all the sites for the day before was higher at 51

millimeters (2.01 inches) than the day the sample was taken at 29 millimeters (1.13 inches). Precipitation grids were computed for both days and the grid generated from the precipitation from 3/3/05 had lower minimum, maximum, and mean statistics for the grid cells over the entire watershed area than the precipitation grid generated for 3/2/05.

The 3/3/05 sampling day appeared to be an outlier for the dataset because it did not belong to the Haberle Road population where less rainfall is generated from the day before the sample is taken. The Dixon-Thompson test was applied to test the sampling point as an outlier for runoff volume. The Dixon-Thompson test is suitable for sample sizes as small as three and can be applied to both low and high outliers (McCuen, 2003). Since, the 3/3/05 data point was the second largest observed runoff volume and the Dixon-Thompson test only tests for the largest or smallest value, the highest runoff volume was not included in the test and a sample size of 11 was used instead of 12. The equation for the Dixon-Thompson High Outlier Test Statistic for sample sizes of 11 is (McCuen, 2003):

$$R = X_{11} - X_9 / X_{11} - X_3 \tag{3.9}$$

where,

X_n = the data are ranked from smallest to largest and the subscript indicates the rank of the value from smallest to largest.

Assuming a normal population, the test statistic (Equation 3.9) was larger than all of the critical values at 5%, 2.5%, and 1% for the sample size of 11. Therefore, the largest runoff volume was rejected and considered an outlier by the Dixon-Thompson test. The point 3/3/05 was then removed from the Haberle Road data set.

3.4.2. Runoff Volume

The model was able to predict the runoff volume at the SH 123 sampling site outlet with very good to unsatisfactory agreement. Adding base flow using the 75% exceedence flow from the FDC developed using SWAT simulated flows resulting in an unsatisfactory agreement for both the E and RSR values (Table 3.4) for all times. Base flow added using the 75% exceedence flow from the FDC developed using instantaneous flows resulted in a very good agreement for both E and RSR values (Table 3.4) across all of the times. The E and RSR values calculated from volumes calculated by adding 100% exceedence flow from an FDC developed using instantaneous flows resulted in a satisfactory to very good agreement. The lag time manually calculated from Alligator Creek resulted in the best performance (Table 3.4) for the baseflow added using exceedence flows from an FDC developed using the instantaneous flows with the time of concentration calculated using SWAT from Alligator Creek performing similarly.

Table 3.4. Model performance for predicting runoff volume at SH 123.

Time	Statistic	75% Exceedence		100% Exceedence
		SWAT	Instantaneous	Instantaneous
Lag Time Manually Calculated from Alligator Creek	E	-1.83	0.95	0.83
	RSR	1.68	0.22	0.41
	RMSE	6217	813	1506
	Observed Average	4496	4496	4496
	Observed Standard Deviation	4128	4128	4128
Time of Concentration Manually Calculated from Alligator Creek	E	-1.11	0.78	0.63
	RSR	1.45	0.47	0.61
	RMSE	8931	2902	4488
	Observed Average	7495	7495	7495
	Observed Standard Deviation	6882	6882	6882
Time of Concentration SWAT Calculated from Alligator Creek	E	-1.95	0.94	0.84
	RSR	1.72	0.25	0.40
	RMSE	6072	872	1405
	Observed Average	4307	4307	4307
	Observed Standard Deviation	3955	3955	3955

The model was able to predict runoff volumes for Haberle Road with very good to unsatisfactory agreement. The travel times calculated manually and with SWAT from Alligator Creek performed with poor to unsatisfactory agreement for E and RSR values (Table 3.5). This implies that there is something happening in the watershed to prevent stream flow from above SH 123 to reach Haberle Road. There was a log jam during the simulation period SH 123 which might have caused stagnation at that site and prohibiting flow from reaching downstream. This may explain why a time of concentration or lag time needs to be calculated from SH 123 to Haberle Road instead of for the entire watershed from Alligator Creek. The performances of the different baseflow methods will only be based upon the times calculated from SH 123 and not

from Alligator Creek. The time of concentration manually calculated from Alligator Creek performed the poorest. Baseflow added using 75% exceedence flow determined from a FDC developed using SWAT simulated flows performed with very good to unsatisfactory agreement for E and RSR values (Table 3.5). The E and RSR values (Table 3.5) calculated using baseflow added using 75% exceedence flow and 100% exceedence flow from an FDC developed using instantaneous flows performed with very good to unsatisfactory agreement.

The lag time SWAT calculated from SH 123 performed the best with very good agreement for E and RSR values (Table 3.5) using all three methods of adding baseflow. The time of concentration SWAT calculated from SH 123 performed with very good to good agreement for all baseflows. The manually calculated lag time performed with fair to unsatisfactory agreement for E and RSR values (Table 3.5). Runoff volumes calculated using the SWAT calculated lag time from SH 123 and the 75% exceedence flow from a FDC developed using instantaneous flows performed the best (Table 3.5).

Table 3.5. Model performance for estimating runoff volume at Haberle Road.

Time	Statistic	75% Exceedence		100% Exceedence
		SWAT	Instantaneous	Instantaneous
Lag Time Manually Calculated from Alligator Creek	E	0.09	-0.05	-0.25
	RSR	0.95	1.03	1.12
	RMSE	21473	23145	25205
	Observed Average	17368	17368	17368
	Observed Standard Deviation	23669	23669	23669
Time of Concentration SWAT Calculated from Alligator Creek	E	0.12	-0.03	-0.22
	RSR	0.94	1.01	1.11
	RMSE	18184	19613	21383
	Observed Average	14881	14881	14881
	Observed Standard Deviation	20279	20279	20279
Lag Time Manually Calculated from SH 123	E	0.35	0.23	0.05
	RSR	0.81	0.88	0.97
	RMSE	5685	6188	6857
	Observed Average	5423	5423	5423
	Observed Standard Deviation	7390	7390	7390
Time of Concentration SWAT Calculated from SH 123	E	0.82	0.77	0.66
	RSR	0.42	0.47	0.59
	RMSE	1011	1133	1401
	Observed Average	1837	1837	1837
	Observed Standard Deviation	2504	2504	2504
Lag Time SWAT Calculated from SH 123	E	0.92	0.95	0.89
	RSR	0.28	0.21	0.33
	RMSE	401	317	476
	Observed Average	1100	1100	1100
	Observed Standard Deviation	1499	1499	1499

The runoff volumes for the SH 123 site were underestimated for three events and overestimated for two events (Figure 3.11 and Figure 3.12). A gap in measuring streamflow and *E. coli* data at SH 123 occurred between the years 2002 and 2010. No data was collected in 2001 where runoff occurred at the same time as a measured *E. coli*

sample. Runoff volumes measured at SH 123 only include the fall and winter seasons with only one sample collected in the spring. This may skew the data distribution because the data do not include the summer season which is typically the driest season for the watershed. Figure 3.11 shows that the data for runoff volumes for SH 123 followed the general trend of the observed runoff volumes and were in close agreement.

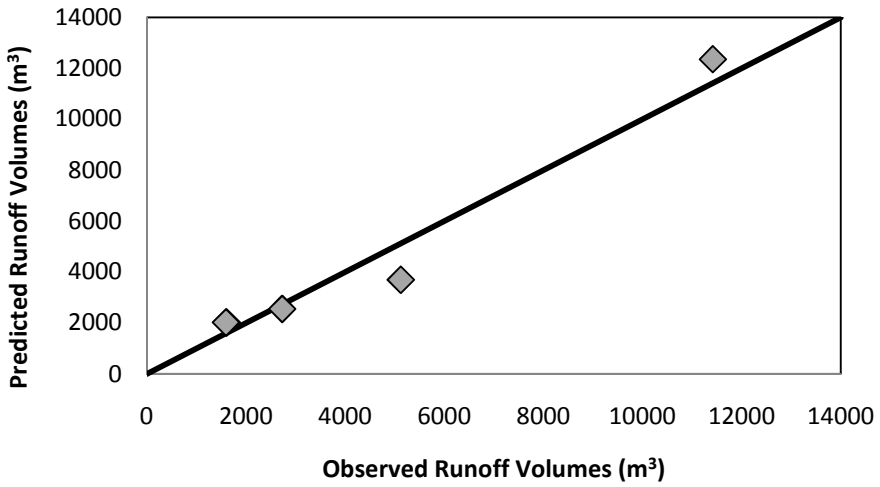

Figure 3.11. Predicted runoff volumes compared to observed runoff volumes using the best performing method for SH 123 for estimating runoff volumes using a time and baseflow.

For the Haberle Road sampling site, simulated runoff volumes were mostly underestimated with the exception of one event shown in Figure 3.12. This data point may have been overestimated because it was taken during the driest season in a year (in August) whereas; the other data were taken in wetter months. The dataset does not include any data points taken in the fall months (October and November); September is

not considered a fall month because the weather is still similar to the summer weather for this region. The dataset also has a gap for the 2009 year where no samples collected had contributing runoff occurring at the same time. This was due to 2009 being an extremely dry year.

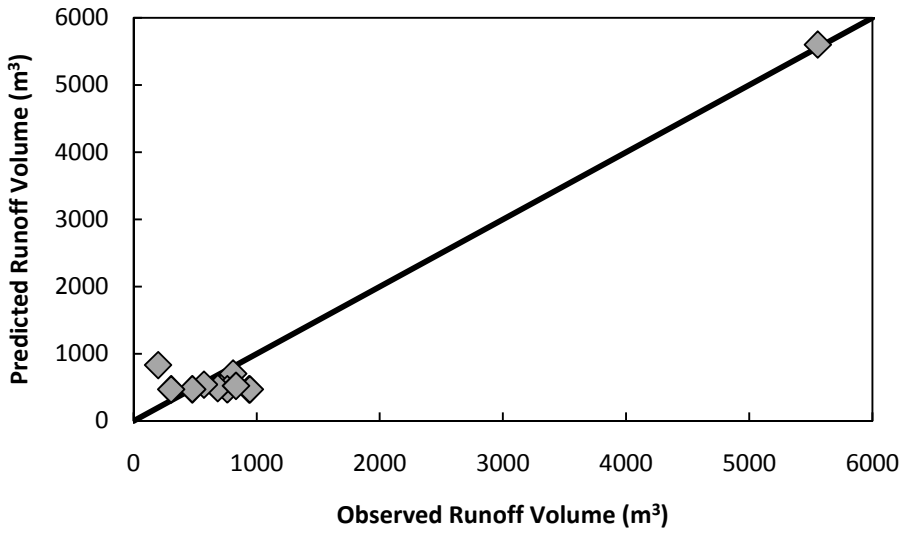

Figure 3.12. Predicted runoff volumes compared to observed runoff volumes for the Haberle Road sampling site using the best performing method for estimating runoff volumes using a time and baseflow.

The spatial watershed model developed from this research was able to predict runoff volumes as well as continuous process models such as SWAT and HSPF were able to predict stream flow. Coffey et al. (2010) were able to validate daily flows using a calibrated SWAT model from January 2004, to February, 2005 with a very good performance rating for *E* of 0.78 in Irish catchments. Parajuli et al. (2009) were able to calibrate the SWAT model in the Upper Wakarusa watershed for mean daily flow of a

subwatershed with very good agreement (E = 0.83) and validate in two subwatersheds belonging to the same watershed as the calibrated subwatershed with very good agreement (E = 0.83 and E = 0.76). SWAT was also run for a watershed with multiple karst features such as multiple springs, sinkholes, and losing streams (Baffaut and Benson, 2009). Baffaut and Benson (2009) validated daily stream flow values from 2001 to 2007 with E values ranging from 0.24 to 0.56 for five stations. Chin et al. (2009) predicted daily and monthly averaged flow for and experimental watershed from 1996 to 2002 using both SWAT and HSPF. SWAT was able to predict monthly and daily flows with very good (E = 0.88) for monthly flows to good (E = 0.65) for daily flows agreement. HSPF predicted both monthly and daily flows with very good agreement (E = 0.89 for monthly and E = 0.87 for daily). Paul et al. (2004) simulated daily mean flow using HSPF in the Salado Creek watershed for a calibration period from 1991 to 1993 with good agreement (E = 0.55). The model from this work estimating runoff volumes from two stations performed similarly to these previously mentioned studies which were chosen because they also predicted *E. coli* or fecal coliform bacteria as well as flow rates, that generally had very good to good agreement using the E values. More data points are needed to calibrate and validate the model properly to show if it is able to consistently and accurately predict runoff volumes. The runoff volumes estimated with the method that performed the best were used to calculate *E. coli* concentrations.

3.4.3. *E. coli* Concentrations

For both the Haberle Road and SH 123 sampling sites, the model predicted *E. coli* concentrations with unsatisfactory agreement (Table 3.6 and Table 3.7) for all four methods of delivery factor calibration for both *E* and *RSR* values. The *RMSE* values for both sites using all four methods, were higher than the observed standard deviations and observed averages (Table 3.6) indicating an unsatisfactory agreement between the observed and predicted *E. coli* concentrations. The delivery factor estimated from the geometric mean of simulated runoff volumes performed the best for Haberle Road while the geometric mean of observed runoff volume performed the best for SH 123. The Haberle Road site consistently performed better than the SH 123 site with the *E* and *RSR* values of -4 and 3 (Table 3.7) for the Haberle Road site and value of -44 and 7 (Table 3.6) for the SH 123 site, respectively, estimated using the best performance values.

Table 3.6. Model performance for *E. coli* concentrations at SH 123.

Concentration Calculation	**Statistic**	**Simulated Delivery Ratio**		**Observed Delivery Ratio**	
		Geomean	**Average**	**Geomean**	**Average**
Observed Flow	E	-90	-441687	-44	-281054
	RSR	10	665	7	530
	RMSE	8	526	5	419
	Observed Average	1.8	1.8	1.8	1.8
	Observed Standard Deviation	0.9	0.9	0.9	0.9
Simulated Flow	E	-142	-656256	-71	-417698
	RSR	12	810	8	646
	RMSE	9	641	7	511
	Observed Average	1.8	1.8	1.8	1.8
	Observed Standard Deviation	0.9	0.9	0.9	0.9

Table 3.7. Model performance for estimating *E. coli* concentrations at Haberle Road.

Concentration Calculation	Statistic	Simulated Delivery Ratio		Observed Delivery Ratio	
		Geomean	Average	Geomean	Average
Observed Flow	E	-56	-1641	-67	-3800
	RSR	8	41	8	62
	RMSE	121	653	133	994
	Observed Average	8	8	8	8
	Observed Standard Deviation	17	17	17	17
Simulated Flow	E	-4	-154	-5	-362
	RSR	2	12	2	19
	RMSE	36	200	40	307
	Observed Average	8	8	8	8
	Observed Standard Deviation	17	17	17	17

The SWAT and HSPF models have been applied to predict *E. coli* concentrations in watersheds with mixed success. Coffey et al. (2010) was able to predict *E. coli* concentrations for Irish catchements using SWAT from grab samples taken monthly from September 2005 to September 2006 resulting in 11 observed samples after removing one outlier. The predicted *E. coli* concentrations were in good agreement with the observed concentrations having an *E* value of 0.59. Parajuli et al. (2009) estimated fecal coliform bacteria concentrations using the SWAT model with unsatisfactory to fair agreement for calibration, validation, and verification watersheds. The calibration watershed was in poor agreement with an *E* value of 0.20 and the validation watershed had an *E* value of 0.31 which resulted in a fair agreement. The verification watershed had an unsatisfactory agreement with an *E* value of -2.2 (Parajuli, et al., 2009). Baffaut and Benson (2009) ran the SWAT model to predict fecal coliform bacteria concentrations calibrated and validated from monthly or bi-monthly grab sample

concentrations for the James River basin which is considered a karst watershed. The SWAT model was calibrated for four different sampling sites having *E* values ranging from -6 to 0.11(Baffaut and Benson, 2009). Chin et al. (2009) predicted fecal coliform bacteria concentrations using both the SWAT and HSPF models for an experimental watershed. SWAT performed better than HSPF with an *E* value of 0.73 compared to an *E* value of 0.33 for HSPF. Paul et al. (2004) did not calibrate the HSPF model due to a lack of observed fecal coliform bacteria data. HSPF was able to simulate in-stream fecal coliform concentrations with good agreement but was unable to capture extreme concentrations (Paul, et al., 2004).

Observed and simulated *E. coli* concentrations predicted using the geometric mean observed and simulated delivery factor for the SH 123 sampling site are presented in Figures 3.13, 3.14, 3.15, and 3.16. The observed *E. coli* concentrations had values ranging from 1.12 to 3.2 CFU/mL. The method of predicting *E. coli* concentrations that had the closest range of concentrations (0.01 to 9.78 CFU/mL) to the observed concentration range was estimated using a delivery factor calculated from the geometric mean of the observed runoff volumes and with the concentration calculated using the observed runoff volumes (Figure 3.15). *E. coli* concentrations predicted using the concentrations calculated from the observed runoff volumes (Figure 3.15 and Figure 3.16) had better results than the concentrations predicted using the concentrations calculated from the simulated runoff volumes (Figure 3.13 and Figure 3.14).

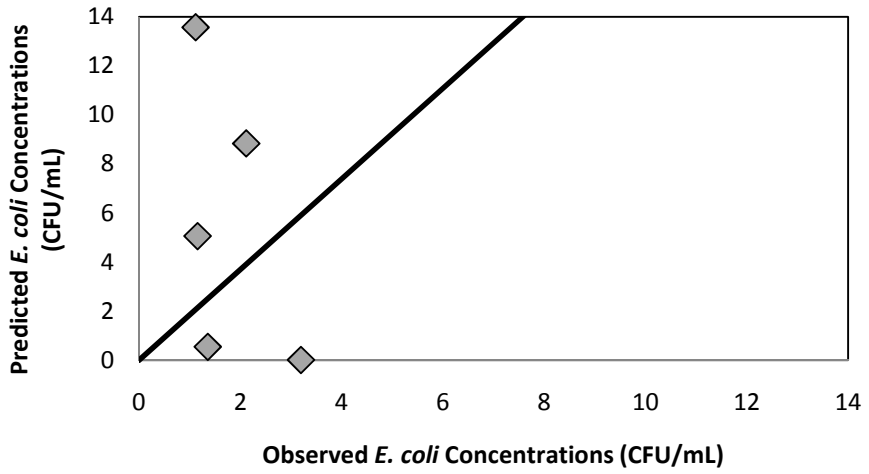

Figure 3.13. Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for SH 123 using instantaneous observed flow delivery factor and the simulated runoff volume.

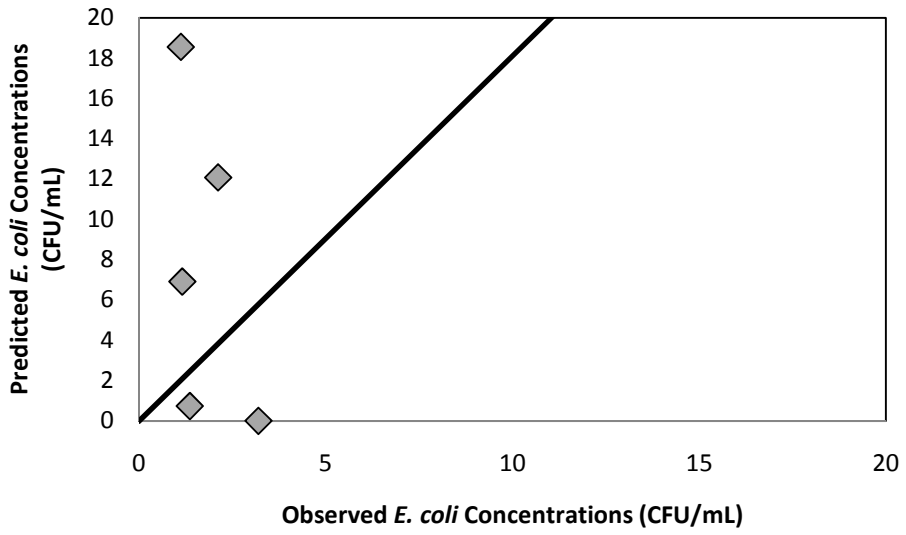

Figure 3.14. Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for SH 123 using simulated flow delivery factor and simulated runoff volume.

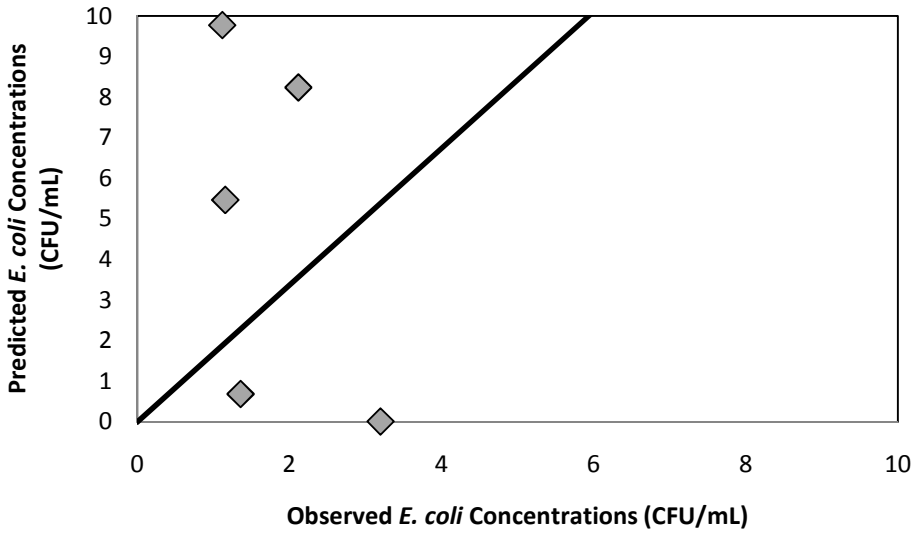

Figure 3.15. Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for SH 123 using instantaneous observed flow delivery factor and observed runoff volume.

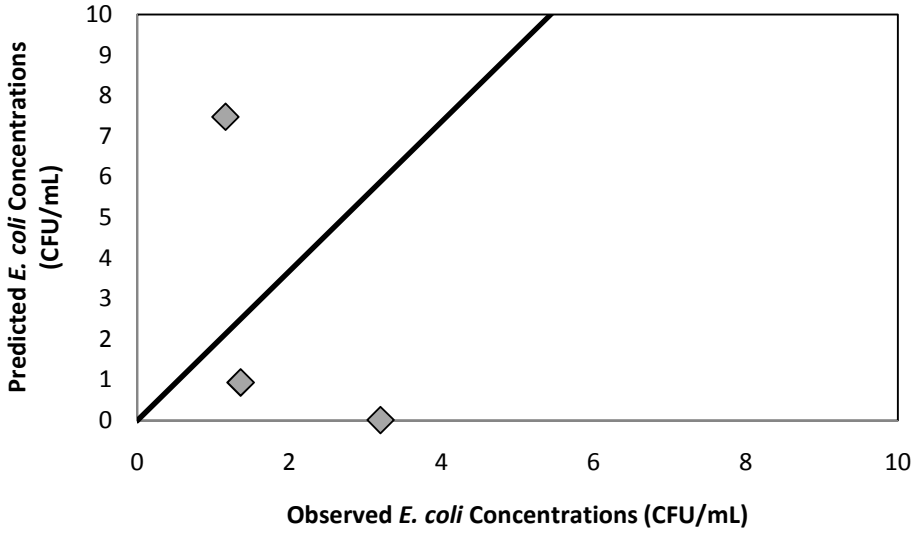

Figure 3.16. Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for SH 123 using simulated flow delivery factor and observed runoff volume.

The prediction of *E. coli* concentrations for the SH 123 sampling site was poorer than the prediction for the Haberle Road sampling site. The SH 123 sampling site followed similar trends as the Haberle Road sampling site. For Haberle Road the delivery factors computed using the simulated runoff volumes performed better than the delivery factors computed using the observed runoff volumes. Concentrations calculated using the simulated runoff volumes performed better than concentrations calculated using the observed runoff volumes. Figures 3.17, 3.18, 3.19, and 3.20 contain the scatter plots comparing the observed and simulated *E. coli* concentrations for Haberle Road using the geometric mean observed and simulated delivery factors. The range for the observed *E. coli* concentrations was from 0.46 to 57 CFU/mL. The *E. coli* concentrations predicted using the delivery factor estimated using the geometric mean of simulated runoff volumes with a concentration calculated using the simulated runoff volumes had the closest range from 0.17 to 96 CFU/mL of predicted concentrations to the observed concentrations. Haberle Road predicted *E. coli* concentrations more accurately than SH 123. This could be because Haberle Road had more samples to calibrate the model.

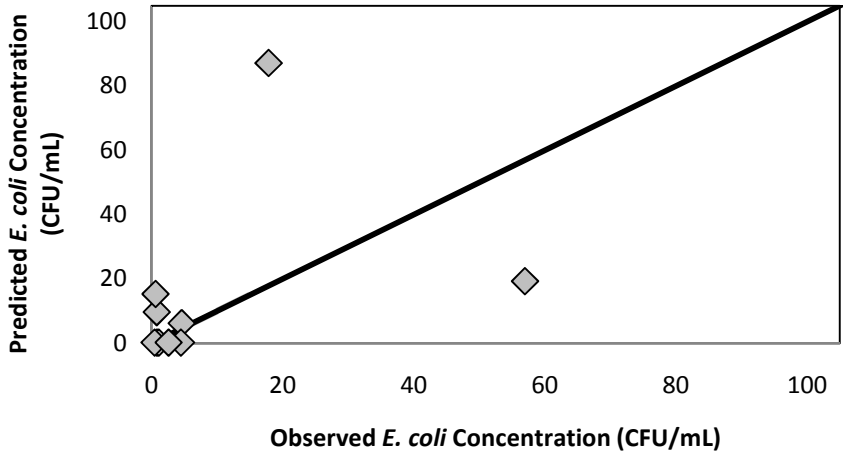

Figure 3.17. Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for Haberle Road using instantaneous observed flow delivery factor and simulated runoff volume.

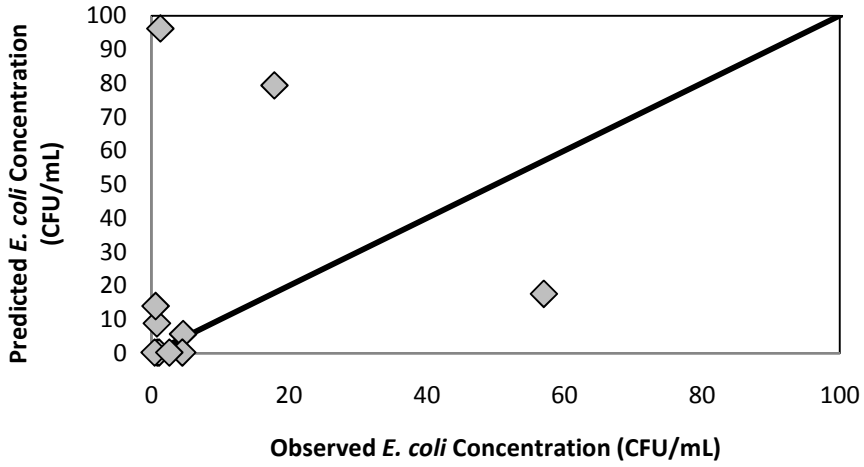

Figure 3.18. Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for Haberle Road using simulated flow delivery factor and simulated runoff volume.

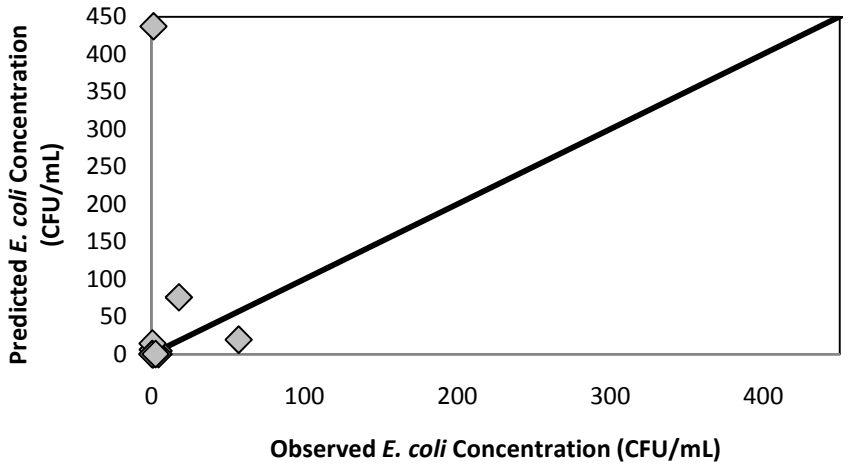

Figure 3.19. Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for Haberle Road using instantaneous observed flow delivery factor and observed runoff volume.

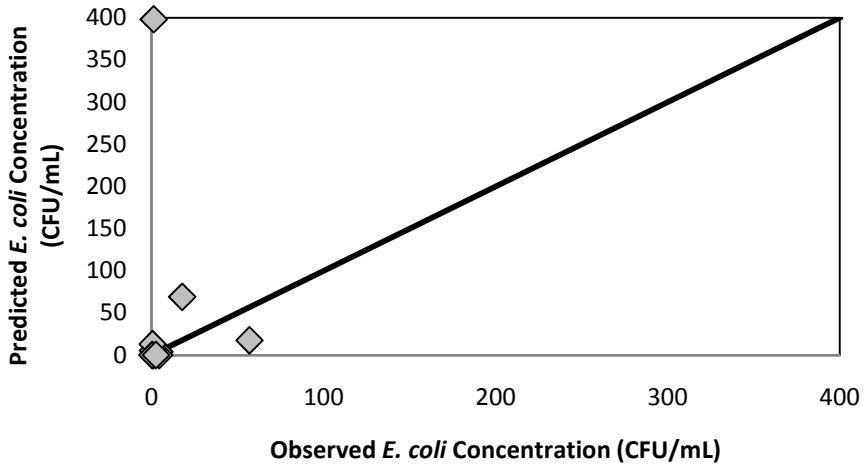

Figure 3.20. Predicted *E. coli* concentrations compared to observed *E. coli* concentrations for Haberle Road using simulated flow delivery factor and observed runoff volume.

3.4.4. Uncertainty

Uncertainty is an important issue regarding water quality modeling because models are increasingly used to guide decisions regarding water resource policy, management, and regulation (Beck, 1987; Sharpley et al., 2002; Harmel et al., 2006; Parajuli et al., 2009). Uncertainty in measured water quality data is introduced during streamflow measurement, sample collection, sample preservation/storage, and laboratory analysis (Harmel et al., 2006). Modeling bacteria transport might have one of the highest probable errors and least confidence compared with modeling surface hydrology, sediment, and nutrients (Novotny, 2003; Parajuli et al., 2009; Coffey et al., 2010). One source of potential uncertainty in the model results can be due to the GIS data inputs (Parajuli et al., 2009). In this study, we used the best available data as inputs into the model including stakeholder input for land use and contributing bacteria source animal numbers and distribution in the watershed. The other GIS inputs including the DEM, soils, and climate data used were the best available data. Harmel et al. (2006) determined the cumulative probable uncertainty for streamflow data ranging from3-42% for best case to worst case scenerios. Uncertainty and variability surrounds bacteria modeling and can lead to large discrepancies in model results (Coffey et al., 2010).

3.4.5. Potential Causes for Inaccurate *E. coli* Bacteria Modeling Results

The assumption that the SELECT *E. coli* load is a constant is the most significant reason for the large discrepancy between the observed and simulated *E. coli* concentrations. The *E. coli* load generated using SELECT is based on the data collected from the stakeholders regarding the densities of the contributing sources as well as the

distribution of those sources for 2010. The densities collected, especially regarding livestock stocking rates and livestock distribution on land use types, vary greatly seasonally and from year to year. The land use and household data determining the distribution of sources is also valid for 2010. One possible reason for the poor performance of the model applied to the SH 123 sampling site is that most of the data for the SH 123 sampling site was collected between 1998 to 2002. On the other hand, the Haberle Road sampling site had data collected from 2004 to 2010, nearer to the time that source data was determined. The *E. coli* loads generated using SELECT for 2010 may be more accurate for the Haberle Road sampling site because less change would have occurred between the earliest sampling date in 2004 and 2010, the date of the source data determination. One solution to improve the model would be to run SELECT for different years with land use and contributing source densities varying from year to year. Accurate data for contributing source densities especially for wildlife and livestock can be difficult to obtain for past years. A compromise to increase accuracy but still obtain accurate data would be to run SELECT for different seasons and vary the land use and contributing source densities for the different seasons.

Another reason for the discrepancies between the observed and predicted *E. coli* concentrations is because the model does not account for point sources contributing *E. coli* directly to the stream. Direct deposition of fecal material by livestock and wildlife is not considered in the model because *E. coli* is considered to be contributing to the stream only through surface runoff. *E. coli* is still being contributed to the stream during low flow conditions. Load Duration Curves (LDCs) were calculated for both the

Haberle Road and SH 123 sampling site using historic data and low flow conditions were determined as a percent exceedence ranging from 75 to 100%. The *E. coli* concentrations for the Haberle and SH 123 sampling sites ranged from 44 to 330 CFU/100 mL and 0 to 438 CFU/100 mL, respectively for both sites for low flow conditions. *E. coli* bacteria occuring in the stream are likely caused by direct depostion or other unknown factors.

The model also does not account for bacteria die-off and re-growth occurring in the stream, soil, and in the fecal material itself. During a rainfall event, sediment located in the stream containing bacteria can be stirred up and further contribute to the *E. coli* concentration occuring in the sample. Coffey et al. (2010) elaborates that there are unknown spatial and temporal sources of contamination contributing bacteria and the ability to accurately account for all of these factors is debateable.

3.5. Conclusions

A watershed model was developed in ArcGIS to estimate the volume of water from runoff and the *E. coli* concentrations contributing at a sampling site. Two sampling sites for the Geronimo Creek watershed were chosen although there was a lack of observed hydrologic and water quality data coinciding with runoff events. Observed streamflow was converted into a runoff volume by removing baseflow and multiplying by the sampling site outlet lag time. A model calibration using four different methods was applied using a delivery factor for the predicted *E. coli* concentrations.

The model results for the runoff volume were in very good agreement ($E = 0.95$, $RSR = 0.21$) for the Haberle Road sampling site and in very good agreement ($E = 0.95$,

RSR = 0.22) for the SH 123 sampling site. The RMSE values were less than half of the standard deviation showing a good agreement between the observed and predicted runoff volumes. The *E. coli* concentration results were in unsatisfactory agreement for both samping sites using all methods. The concentrations calculated with the geometric mean delivery factor performed the best for both sites. The Haberle Road sampling site performed consistently better for all methods than the SH 123 sampling site.

The model was unable to accurately predict the *E. coli* concentrations occuring in stream. One potential reason for the model inaccuracies for predicting *E. coli* concentrations is applying the *E. coli* load resulting from SELECT as a constant. This may be overcome by varying the SELECT *E. coli* loads for different years and/or seasons. Direct deposition is not considered in the model although *E. coli* is contributing into the stream during periods where there are low flow conditions. Bacteria die-off and re-growth occuring in the stream, soil,and fecal material is not considered in the model and if it were it could potentially increase the accuracy of the model at predicting *E. coli* concentrations. There are other unknown factors that contribute to *E. coli* bacteria contamination in streams which make predicting *E. coli* concentrations with a model difficult. Although the model did not accurately predict *E. coli* concentrations, it can be modified in multiple ways to increase the model accuracy by varying the *E. coli* concentration yearly and seasonally, accounting for direct deposition, and accounting for die-off and regrowth.

CHAPTER IV

CONCLUSIONS

4.1. Conclusions

1. The SELECT methodology was applied to three watersheds Buck Creek, the five tributaries of the Little Brazos River, and Lampasas River to support the development of watershed protection plans (WPPs) and can easily be adapted to different watersheds and reflect the potential sources, stakeholder concerns, and data availability of the watershed.
2. An automated watershed model was developed to convert the *E. coli* loads resulting from SELECT analysis into an *E. coli* concentration occurring in the stream. The tool was automatically able to calculate runoff volumes resulting at a drainage area outlet for a rainfall event.
3. The runoff volumes were predicted with very good to good agreement for both the SH 123 and Haberle Road sampling sites in the Geronimo Creek watershed.
4. *E. coli* concentrations were predicted with unsatisfactory agreement for both the SH 123 and Haberle Road sampling sites using four different methods of delivery factor calibration.

4.2. Recommendations

The model can be improved by not appling the SELECT *E. coli* loads as a constant developed from the most recent data collected for land use and contributing sources. Land use and contributing source densities should be collected for individual years and seasons and the SELECT methodology should be applied for the differing

years and seasons. The SELECT *E. coli* loads input into the model should be varied for different years and/or seasons to better account of the changes influencing contributing sources. A baseflow *E. coli* concentration should be determined or a method to calculate the *E. coli* contribution into the stream due to direct deposition should considered and incorporated into the model. Bacteria die-off and re-growth occuring in the stream, soil,and fecal material should be explored and then incorporated into the model to better represent the fate and transport of bacteria into the stream. Other unknown factors that contribute to *E. coli* bacteria contamination in streams should be considered or accounted for in the model through an additional model parameter. Although the model did not accurately predict *E. coli* concentrations, it can be modified in multiple ways to improve the model predictions.

REFERENCES

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R., 1998. Large area hydrologic modeling and assessment part I: model development. Journal of the American Water Resources Association. 34 (1), 73-89.

Baffaut, C., & Benson, V. W., 2009. Modeling flow and pollutant transport in a karst watershed with SWAT. Transactions of the ASABE. 52 (2), 469-479.

Beck, M. B., 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resources Research. 23 (8), 1393-1442.

Benham, B., Baffault, C., Zeckoski, R., Mankin, K., Pachepsky, Y., Sedeghi, A., et al., 2006. Modeling bacteria fate and transport in watersheds to support TMDLs. Transactions of the ASABE. 49 (4), 987-1002.

Bicknell, B., Imoff, J., Kittle, J. J., & Donigian, J. A., 2001. Hydrological simulation program - fortran (HSPF): user's manual for release 12. Athens, GA: U.S. Environmental Protection Agency, National Exposure Research Laboratory.

Bora, D. K., & Bera, M., 2004. Watershed-scale hydrologic and nonpoint-source pollution models: review of applications. Transactions of the ASABE. 47 (3), 789-803.

Chin, D. A., Sakura-Lemessy, D., Bosch, D. D., & Gay, P. A., 2009. Watershed-scale fate and transport of bacteria. Transactions of the ASABE. 52 (1), 145-154.

Coffey, R., Cummins, E., Bhreathnach, N., Flaherty, V. O., & Cormican, M., 2010. Development of a pathogen transport model for Irish catchments using SWAT. Agricultural Water Management. 97, 101-111.

Curriero, F. C., Patz, J. A., Rose, J. B., & Lele, S., 2001. The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948-1994. American Journal of Public Health. 97 (8), 1194-1199.

Dictson, N., 2009. Geronimo and Alligator Creeks watershed partnership information sheet. Retrieved August 13, 2010, from Geronimo and Alligator Creeks watershed partnership: http://www.geronimocreek.org/GeronimoInfoSheet.pdf

Fraser, R. H., Barten, P. K., & Pinney, D. A., 1998. Predicting stream pathogen loading from livestock using a geographical information system-based delivery model. Journal of Environmental Quality. 27 , 935-945.

Haan, C., Barfield, B., & Hayes, J., 1994. Design Hydrology and Sedimentology for Small Catchments. San Diego, CA: Academic Press, Inc.

Harmel, R. D., Cooper, R. J., Slade, R. M., Haney, R. L., & Arnold, J. G., 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. Transactions of the ASABE. 49 (3), 689-701.

Legates, D. R., & McCabe, G. J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resources Research. 35 (1), 233-241.

Lockwood, M., 2005. White-tailed deer population trends. Austin, TX: Texas Parks and Wildlife Department.

McCuen, R. H., 2003. Modeling Hydrologic Change: Statistical Methods. Boca Raton, FL: CRC.

McCuen, R. H., Knight, Z., & Cutter, A. G., 2006. Evaluation of the Nash-Sutcliffe efficiency index. Journal of Hydrologic Engineering. 11 (6), 597-602.

McElroy, A. D., Chiu, S. Y., Nebgen, J. W., Aleti, A., & Bennett, F. W., 1976. Loading functions for assessment of water pollution from nonpoint sources. Washington, DC: U.S. Environmental Protection Agency. EPA-600/2-76-151. USEPA Office of Resources and Development.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L., 2007. Model evaluation guidelines for systematic quantifications of accuracy in watershed simulations. Transactions of the ASABE. 50 (3), 885-900.

Munafo, M., Cecchi, G., Baiocco, F., & Mancini, L., 2005. River pollution from non-point sources: a new simplified method of assessment. Journal of Environmental Management. 77, 93-98.

Nash, J. E., & Sutcliffe, I. V., 1970. River flow forecasting trhough conceptual models part 1 - a discussion of principles. Journal of Hydrology. 10, 282-290.

Novotny, V. (2003). Water Quality, Diffuse Pollution and Watershed Management. Hoboken, NJ: John Wiley and Sons Inc.

Parajuli, P. B., Mankin, K. R., & Barnes, P. L., 2009. Source specific fecal bacteria modeling using soila and water assessment tool model. Biosource Technology. 100, 953-963.

Paul, S., Haan, P. K., Matlock, M. D., Mukhtar, S., & Pillai, S. D., 2004. Analysis of the HSPF water quality parameter unvertainty in predicting peak in-stream feacal coliform concentrations. Transactions of the ASAE. 47 (1), 69-78.

Riebschleager, K. J., Karthikeyan, R., & Srinivasan, R., 2011. Estimating potential *E. coli* sources in a watershed using spatially explicit modeling techniques. (Under review).

Sadeghi, A. M., & Arnold, J. G., 2002. A SWAT/microbial sub-model for predicting pathogent loadings in surface and groundwater at watershed and basin scales. total maximum daily load (TMDL) environmental regulations. St. Joseph, MI: ASAE.

Schuurmans, J. M., & Bierkens, M. F., 2006. Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model. Hydrology and Earth System Sciences. 3, 677-693.

Sharpley, A. N., Kleinman, P. J., McDowell, R. W., Gitau, M., & Bryant, R. B., 2002. Modeling phosphorus transport in agricultural watersheds: processes and possibilities. Journal of Soil & Water Conservation. 57 (6), 425-439.

Singh, J., Knapp, H. V., & Demissie, M., 2004. Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. ISWS CR 2004-08. Champaign, IL: Illinois State Water Survey.

Soil Conservation Service, 1986. Urban hydrology for small watersheds, Technical Release No. 55. Washington, DC: Soil Conservation Service, U.S. Department of Agriculture.

TCEQ, 2008. 2008 Texas water quality inventory and 303(d) list. Retrieved August 4, 2010, from Texas Commission on Environmental Quality:

http://www.tceq.state.tx.us/compliance/monitoring/water/quality/data/08twqi/twqt08.html.

Teague, A., Karthikeyan, R., Babbar-Sebens, M., Srinivasan, R., & Persyn, R. A., 2009. Spatially explicit load enrichment calculation tool to identify potential *E. coli* sources in watersheds. Transactions of the ASABE. 52 (4), 1109-1120.

USDA-NASS, 2007. 2007 census of agriculture: county data. Washington, DC: USDA National Agricultural Statistics Survey.

USEPA, 2001a. Protocol for developing pathogen TMDLs: source assessment, 5-1 - 5-18 1st ed. EPA841-R-00-002. Washington, DC: U.S. Environmental Protection Agency, Office of Water.

USEPA, 2001b. The national costs of the total maximum daily load program (draft report), EPA 841-D-01-003. Washington DC: U.S. Environmental Protection Agency, Office of Water.

USEPA, 2003. Bacterial water quality standards for recreational waters (freshwater and marine waters) status report, EPA 823-R-03-008. Washington, DC: U.S. Environmental Protection Agency, Office of Water.

USEPA, 2008. Causes of impairment for 303(d) listed waters. Retrieved April 24, 2011, from U.S. Environmental Protection Agency: http://iaspub.epa.gov/waters10/attains_nation_cy.control?p_report_type=T.

USGS, 2010. HSPF Hydrological simulation program--Fortran. Retrieved August 5, 2010, from U.S. Geological Survey: http://water.usgs.gov/software/HSPF/

Zhang, X., & Srinivasan, R., 2009. GIS-based spatial precipitation estimation: a comparison of geostatistical approaches. Journal of the American Water Resources Association. 45(4) , 894-906.

VITA

Name:	Kyna Ellen McKee
Address:	Department of Biological and Agricultural Engineering c/o Dr. Karthikeyan Texas A&M University College Station, TX 77840-2117
Email Address:	kyna.mckee@gmail.com
Education:	B.S., Agricultural Engineering, Texas A&M University, 2009 M.S., Biological and Agricultural Engineering Texas A&M University, 2011